Supplemental Material to
Regression-Adjusted Estimation of Quantile Treatment Effects under Covariate-Adaptive Randomizations

By

Liang Jiang, Peter C.B. Phillips, Yubo Tao, and Yichong Zhang

May 2021

COWLES FOUNDATION DISCUSSION PAPER NO. 2288S

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Supplement to "Regression-Adjusted Estimation of Quantile Treatment Effects under Covariate-Adaptive Randomizations"*

Liang Jiang[†] Peter C.B. Phillips [‡] Yubo Tao[§] Yichong Zhang[¶]

May 31, 2021

Abstract

This document collects supplementary materials for the main paper. Section A presents additional simulation results for pointwise tests at 25% and 75% quantiles. Section B reports addition simulation results when the true, instead of the estimated, propensity score is used. Section C derives the pseudo true value of the logistic regression that minimizes the asymptotic variance of the difference of two quantile treatment effect (QTE) estimators. Section D proposes a way to consistently estimate one point in the set of pseudo true values that achieves the minimum of the QTE estimator's asymptotic variance. Section E introduces extra notation. Sections F–O prove Theorems 3.1, 4.1, 5.1, 5.2, Propositions 5.1, 5.2, 5.3, Theorems 5.4, 5.5, and Proposition D.1, respectively. Section P collects all the supporting technical lemmas.

Keywords: Covariate-adaptive randomization, high-dimensional data, regression adjustment

JEL codes: C14, C21, I21

A Additional Simulation Results for Pointwise Tests

This section gives additional simulation results for pointwise tests at 25% and 75% quantiles. The results are summarized in Tables A1 and A2. The simulation settings are the same as the pointwise test simulations in Section 7 of the original paper.

*Yichong Zhang acknowledges the financial support from Singapore Ministry of Education Tier 2 grant under grant MOE2018-T2-2-169 and a Lee Kong Chian fellowship. Phillips acknowledges support from NSF Grant No. SES 18-50860, a Kelly Fellowship at the University of Auckland, and a Lee Kong Chian Fellowship. Any errors are our own.

[†]Fudan University. E-mail address: jiangliang@fudan.edu.cn.

[‡]Yale University, University of Auckland, University of Southampton, and Singapore Management University. E-mail address: peter.phillips@yale.edu

[§]Singapore Management University. E-mail address: ybtao@smu.edu.sg.

[¶]Singapore Management University. E-mail address: yczhang@smu.edu.sg.

Table A1: Pointwise Test ($\tau = 0.25$)

	$N = 200$				$N = 400$			
Methods	SRS	WEI	BCD	SBR	SRS	WEI	BCD	SBR
				Panel A: DGP (i)				
A.1: Size								
NA	0.055	0.052	0.051	0.051	0.054	0.054	0.053	0.053
LP	0.055	0.054	0.058	0.054	0.055	0.053	0.052	0.052
LG	0.052	0.048	0.054	0.051	0.055	0.051	0.053	0.053
ML	0.060	0.059	0.060	0.059	0.052	0.053	0.055	0.056
NP	0.063	0.064	0.060	0.061	0.053	0.058	0.061	0.060
A.2: Power								
NA	0.341	0.329	0.346	0.352	0.591	0.593	0.589	0.611
LP	0.404	0.406	0.402	0.397	0.688	0.696	0.691	0.695
LG	0.386	0.392	0.399	0.390	0.692	0.684	0.688	0.697
ML	0.435	0.423	0.428	0.426	0.718	0.729	0.733	0.724
NP	0.443	0.437	0.434	0.443	0.730	0.723	0.737	0.728
				Panel B: DGP (ii)				
B.1: Size								
NA	0.047	0.050	0.048	0.047	0.046	0.049	0.048	0.048
LP	0.052	0.050	0.049	0.055	0.049	0.047	0.052	0.049
LG	0.046	0.047	0.040	0.041	0.049	0.042	0.047	0.042
ML	0.063	0.062	0.057	0.065	0.046	0.051	0.049	0.051
NP	0.066	0.067	0.064	0.063	0.058	0.057	0.048	0.052
B.2: Power								
NA	0.447	0.457	0.460	0.487	0.741	0.738	0.740	0.759
LP	0.528	0.523	0.539	0.538	0.812	0.822	0.819	0.819
LG	0.462	0.465	0.467	0.467	0.789	0.790	0.791	0.796
ML	0.569	0.578	0.564	0.575	0.841	0.846	0.844	0.838
NP	0.576	0.567	0.574	0.572	0.843	0.843	0.843	0.845

Table A2: Pointwise Test ($\tau = 0.75$)

	$N = 200$				$N = 400$			
Methods	SRS	WEI	BCD	SBR	SRS	WEI	BCD	SBR
				Panel A: DGP (i)				
A.1: Size								
NA	0.055	0.056	0.058	0.050	0.054	0.055	0.051	0.056
LP	0.053	0.053	0.049	0.050	0.050	0.050	0.053	0.052
LG	0.065	0.071	0.069	0.069	0.065	0.067	0.065	0.065
ML	0.052	0.060	0.058	0.061	0.058	0.057	0.058	0.056
NP	0.062	0.067	0.063	0.063	0.059	0.054	0.057	0.056
A.2: Power								
NA	0.337	0.341	0.343	0.332	0.583	0.594	0.590	0.571
LP	0.426	0.434	0.430	0.441	0.685	0.694	0.696	0.698
LG	0.415	0.416	0.415	0.414	0.662	0.668	0.666	0.671
ML	0.425	0.430	0.424	0.426	0.697	0.708	0.709	0.708
NP	0.438	0.437	0.433	0.443	0.714	0.717	0.714	0.704
				Panel B: DGP (ii)				
B.1: Size								
NA	0.051	0.053	0.048	0.047	0.055	0.052	0.053	0.051
LP	0.057	0.054	0.053	0.055	0.052	0.056	0.055	0.055
LG	0.073	0.076	0.072	0.075	0.070	0.070	0.073	0.066
ML	0.062	0.068	0.063	0.064	0.058	0.055	0.056	0.059
NP	0.069	0.070	0.065	0.067	0.058	0.062	0.058	0.056
B.2: Power								
NA	0.318	0.334	0.332	0.308	0.550	0.551	0.554	0.531
LP	0.383	0.393	0.398	0.391	0.633	0.629	0.637	0.621
LG	0.405	0.404	0.409	0.403	0.633	0.632	0.639	0.622
ML	0.411	0.428	0.416	0.427	0.658	0.672	0.666	0.663
NP	0.413	0.417	0.416	0.417	0.662	0.660	0.657	0.671

B Additional Simulation Results for Tests with Naïve Estimator

Table A3: Pointwise Test with Naïve Estimator ($\tau = 0.25$, $\hat{\pi}(s) = 0.5$)

	$N = 200$				$N = 400$			
Methods	SRS	WEI	BCD	SBR	SRS	WEI	BCD	SBR
				Panel A: DGP (i)				
A.1: Size								
NA	0.051	0.029	0.023	0.021	0.052	0.030	0.022	0.022
LP	0.048	0.019	0.010	0.007	0.054	0.016	0.007	0.006
LG	0.048	0.029	0.023	0.023	0.052	0.039	0.033	0.033
ML	0.050	0.033	0.033	0.032	0.051	0.045	0.037	0.040
NP	0.057	0.040	0.038	0.038	0.057	0.049	0.047	0.040
A.2: Power								
NA	0.253	0.235	0.217	0.237	0.465	0.447	0.441	0.459
LP	0.222	0.177	0.141	0.125	0.410	0.379	0.362	0.327
LG	0.313	0.285	0.264	0.251	0.608	0.605	0.597	0.597
ML	0.367	0.339	0.314	0.298	0.683	0.665	0.662	0.638
NP	0.394	0.356	0.344	0.323	0.690	0.670	0.673	0.661
				Panel B: DGP (ii)				
B.1: Size								
NA	0.051	0.021	0.009	0.008	0.049	0.020	0.011	0.009
LP	0.048	0.011	0.002	0.002	0.051	0.012	0.002	0.002
LG	0.037	0.025	0.018	0.016	0.039	0.024	0.018	0.019
ML	0.049	0.045	0.048	0.049	0.046	0.049	0.050	0.050
NP	0.061	0.056	0.058	0.060	0.055	0.053	0.053	0.048
B.2: Power								
NA	0.283	0.244	0.228	0.246	0.503	0.510	0.505	0.525
LP	0.234	0.193	0.155	0.119	0.439	0.402	0.397	0.368
LG	0.342	0.323	0.309	0.322	0.647	0.655	0.665	0.666
ML	0.529	0.530	0.535	0.531	0.826	0.830	0.830	0.833
NP	0.557	0.546	0.557	0.562	0.840	0.843	0.845	0.843

Table A4: Pointwise Test with Naïve Estimator ($\tau = 0.75$, $\hat{\pi}(s) = 0.5$)

	$N = 200$				$N = 400$			
Methods	SRS	WEI	BCD	SBR	SRS	WEI	BCD	SBR
Panel A: DGP (i)								
A.1: Size								
NA	0.054	0.034	0.023	0.025	0.054	0.034	0.018	0.023
LP	0.017	0.001	0.000	0.000	0.029	0.003	0.000	0.000
LG	0.040	0.006	0.001	0.001	0.049	0.006	0.000	0.001
ML	0.030	0.003	0.000	0.001	0.041	0.006	0.000	0.001
NP	0.031	0.004	0.000	0.000	0.043	0.006	0.001	0.000
A.2: Power								
NA	0.259	0.243	0.237	0.217	0.461	0.455	0.453	0.439
LP	0.090	0.025	0.002	0.000	0.157	0.071	0.012	0.005
LG	0.132	0.067	0.024	0.012	0.221	0.135	0.071	0.056
ML	0.124	0.049	0.014	0.007	0.225	0.137	0.072	0.049
NP	0.125	0.057	0.016	0.011	0.231	0.142	0.067	0.053
Panel B: DGP (ii)								
B.1: Size								
NA	0.053	0.028	0.018	0.017	0.049	0.026	0.017	0.017
LP	0.010	0.000	0.000	0.000	0.023	0.002	0.000	0.000
LG	0.041	0.005	0.000	0.000	0.051	0.006	0.000	0.000
ML	0.024	0.003	0.001	0.000	0.042	0.004	0.000	0.000
NP	0.025	0.003	0.001	0.000	0.044	0.004	0.000	0.000
B.2: Power								
NA	0.232	0.215	0.191	0.176	0.405	0.389	0.389	0.373
LP	0.066	0.020	0.001	0.000	0.144	0.055	0.008	0.003
LG	0.121	0.052	0.017	0.010	0.207	0.122	0.053	0.039
ML	0.105	0.040	0.009	0.004	0.197	0.113	0.041	0.029
NP	0.106	0.044	0.011	0.005	0.196	0.116	0.048	0.032

Table A5: Test for Differences with Naïve Estimator ($\tau_1 = 0.25$, $\tau_2 = 0.75$, $\hat{\pi}(s) = 0.5$)

	$N = 200$				$N = 400$			
Methods	SRS	WEI	BCD	SBR	SRS	WEI	BCD	SBR
				Panel A: DGP (i)				
A.1: Size								
NA	0.041	0.034	0.028	0.027	0.041	0.032	0.029	0.033
LP	0.010	0.007	0.003	0.002	0.031	0.010	0.003	0.003
LG	0.030	0.006	0.001	0.002	0.034	0.008	0.002	0.002
ML	0.027	0.008	0.002	0.002	0.038	0.010	0.002	0.002
NP	0.028	0.007	0.002	0.002	0.043	0.010	0.003	0.003
A.2: Power								
NA	0.187	0.179	0.171	0.157	0.344	0.341	0.343	0.311
LP	0.082	0.058	0.037	0.027	0.197	0.150	0.110	0.093
LG	0.095	0.048	0.027	0.020	0.172	0.113	0.068	0.056
ML	0.106	0.059	0.031	0.023	0.191	0.127	0.079	0.066
NP	0.103	0.058	0.030	0.021	0.187	0.126	0.078	0.066
				Panel B: DGP (ii)				
B.1: Size								
NA	0.037	0.029	0.022	0.023	0.042	0.026	0.023	0.024
LP	0.007	0.003	0.000	0.001	0.023	0.005	0.001	0.001
LG	0.024	0.003	0.001	0.000	0.037	0.005	0.000	0.000
ML	0.021	0.005	0.000	0.000	0.039	0.005	0.001	0.001
NP	0.023	0.004	0.001	0.000	0.035	0.007	0.001	0.000
B.2: Power								
NA	0.175	0.151	0.153	0.125	0.323	0.312	0.314	0.283
LP	0.056	0.036	0.016	0.008	0.168	0.108	0.064	0.046
LG	0.073	0.028	0.009	0.004	0.139	0.066	0.030	0.019
ML	0.082	0.033	0.010	0.006	0.177	0.093	0.042	0.033
NP	0.085	0.037	0.014	0.007	0.168	0.097	0.044	0.037

Table A6: Uniform Test with Naïve Estimator ($\tau \in [0.25, 0.75]$, $\hat{\pi}(s) = 0.5$)

	$N = 200$				$N = 400$			
Methods	SRS	WEI	BCD	SBR	SRS	WEI	BCD	SBR
				Panel A: DGP (i)				
A.1: Size								
NA	0.045	0.020	0.011	0.012	0.049	0.022	0.014	0.012
LP	0.031	0.004	0.001	0.001	0.036	0.005	0.001	0.001
LG	0.037	0.010	0.005	0.004	0.049	0.015	0.007	0.006
ML	0.036	0.011	0.007	0.008	0.044	0.015	0.011	0.011
NP	0.038	0.014	0.010	0.010	0.047	0.016	0.013	0.011
A.2: Power								
NA	0.298	0.269	0.250	0.244	0.562	0.560	0.577	0.564
LP	0.174	0.090	0.059	0.039	0.350	0.291	0.231	0.188
LG	0.282	0.210	0.174	0.146	0.602	0.554	0.536	0.505
ML	0.320	0.248	0.207	0.184	0.668	0.642	0.620	0.592
NP	0.342	0.280	0.238	0.206	0.682	0.649	0.630	0.609
				Panel B: DGP (ii)				
B.1: Size								
NA	0.043	0.012	0.005	0.004	0.047	0.015	0.004	0.005
LP	0.028	0.001	0.000	0.000	0.041	0.003	0.000	0.000
LG	0.036	0.009	0.004	0.004	0.045	0.009	0.004	0.005
ML	0.033	0.016	0.013	0.017	0.042	0.024	0.023	0.022
NP	0.041	0.023	0.024	0.023	0.047	0.026	0.032	0.027
B.2: Power								
NA	0.302	0.258	0.221	0.219	0.562	0.574	0.577	0.572
LP	0.167	0.087	0.042	0.025	0.343	0.282	0.215	0.182
LG	0.337	0.258	0.228	0.221	0.688	0.653	0.626	0.623
ML	0.532	0.503	0.482	0.471	0.896	0.886	0.880	0.880
NP	0.591	0.552	0.535	0.526	0.918	0.910	0.896	0.905

C The Optimal Pseudo True Value when Inferring the Difference of Two QTEs

The pseudo true value $\theta^{LG}_{a,s}(\tau)$ is defined to achieve the minimum asymptotic variance of $\hat{q}^{par}(\tau)$ under the logistic model. However, it does not necessarily minimize the asymptotic variance of $\hat{q}^{par}(\tau_1) - \hat{q}^{par}(\tau_2)$, which is used to construct the test statistic for the second null hypothesis in Section 6.2.[1] In this section we derive the pseudo true value that minimizes the asymptotic variance of $\hat{q}^{par}(\tau_1) - \hat{q}^{par}(\tau_2)$ and its estimator for the logistic model. Proof of the consistency of the estimator and verification of Assumptions 3 and 5 are similar to those for $\hat{\theta}^{LP}_{a,s}(\tau)$ and $\hat{\theta}^{LG}_{a,s}(\tau)$

[1] For the linear probability model Theorem 5.2 has shown that $(\theta^{LG}_{a,s}(\tau_1), \theta^{LP}_{a,s}(\tau_2))_{a=0,1,s\in\mathcal{S}}$ still minimizes the asymptotic variance of $\hat{q}^{par}(\tau_1) - \hat{q}^{par}(\tau_2)$.

studied in Section 5.1 and are therefore omitted for brevity. We first state a general result that is parallel to Theorem 5.2.

Theorem C.1. *Suppose Assumptions 1, 2, 4, 6 hold, and $\Lambda_{a,s}(X_i, \theta_{a,s}(\tau_1,\tau_2))$ is differentiable in $\theta_{a,s}(\tau_1,\tau_2)$. Then, the asymptotic variance of $\hat{q}^{par}(\tau_1) - \hat{q}^{par}(\tau_2)$ is minimized at*

$$(\theta_{1,s}(\tau_1), \theta_{0,s}(\tau_1), \theta_{1,s}(\tau_2), \theta_{0,s}(\tau_2)) \in \Theta_s(\tau_1,\tau_2),$$

where for $s \in \mathcal{S}$ and $\tau_1, \tau_2, \in \Upsilon$,

$$\begin{aligned}
\Theta_s(\tau_1,\tau_2) = \mathop{\arg\min}_{\theta_{1,1},\theta_{0,1},\theta_{1,2},\theta_{0,2}} \quad & Q(s,\tau_1,\theta_{1,1},\theta_{0,1}) + Q(s,\tau_2,\theta_{1,2},\theta_{0,2}) - 2\tilde{Q}(s,\tau_1,\tau_2,\theta_{1,1},\theta_{0,1},\theta_{1,2},\theta_{0,2}), \\
& \tilde{Q}(s,\tau_1,\tau_2,\theta_{1,1},\theta_{0,1},\theta_{1,2},\theta_{0,2}) \qquad \text{(C.1)} \\
=& \mathbb{E}\Bigg\{ \left(\frac{g_{1,s}(X_i,\theta_{1,1})}{f_1(q_1(\tau_1))} + \frac{\pi(s)}{1-\pi(s)} \frac{g_{0,s}(X_i,\theta_{0,1})}{f_0(q_0(\tau_1))} \right) \left(\frac{g_{1,s}(X_i,\theta_{1,2})}{f_1(q_1(\tau_2))} + \frac{\pi(s)}{1-\pi(s)} \frac{g_{0,s}(X_i,\theta_{0,2})}{f_0(q_0(\tau_2))} \right) \\
& - \left(\frac{g_{1,s}(X_i,\theta_{1,1})}{f_1(q_1(\tau_1))} + \frac{\pi(s)}{1-\pi(s)} \frac{g_{0,s}(X_i,\theta_{0,1})}{f_0(q_0(\tau_1))} \right) \left(\frac{m_1(\tau_2,s,X_i) - m_1(\tau_2,s)}{f_1(q_1(\tau_2))} \right) \\
& - \left(\frac{g_{1,s}(X_i,\theta_{1,2})}{f_1(q_1(\tau_2))} + \frac{\pi(s)}{1-\pi(s)} \frac{g_{0,s}(X_i,\theta_{0,2})}{f_0(q_0(\tau_2))} \right) \left(\frac{m_1(\tau_1,s,X_i) - m_1(\tau_1,s)}{f_1(q_1(\tau_1))} \right) \\
& - \frac{\pi(s)}{1-\pi(s)} \left(\frac{g_{1,s}(X_i,\theta_{1,1})}{f_1(q_1(\tau_1))} + \frac{\pi(s)}{1-\pi(s)} \frac{g_{0,s}(X_i,\theta_{0,1})}{f_0(q_0(\tau_1))} \right) \left(\frac{m_0(\tau_2,s,X_i) - m_0(\tau_2,s)}{f_0(q_0(\tau_2))} \right) \\
& - \frac{\pi(s)}{1-\pi(s)} \left(\frac{g_{1,s}(X_i,\theta_{1,2})}{f_1(q_1(\tau_2))} + \frac{\pi(s)}{1-\pi(s)} \frac{g_{0,s}(X_i,\theta_{0,2})}{f_0(q_0(\tau_2))} \right) \left(\frac{m_0(\tau_1,s,X_i) - m_0(\tau_1,s)}{f_0(q_0(\tau_1))} \right) \Bigg| S_i = s \Bigg\}, \qquad \text{(C.2)}
\end{aligned}$$

$Q(s,\tau,\theta_1,\theta_0)$ is defined in (5.3), and $g_{a,s}(X_i,\theta_a) = \mathbb{E}(\Lambda_{a,s}(X_i,\theta_a)|S_i = s) - \Lambda_{a,s}(X_i,\theta_a)$.

For the logistic model, we need to assume $\Theta_s(\tau_1,\tau_2)$ is a singleton, denoted as

$$(\theta^{LG}_{1,s}(\tau_1), \theta^{LG}_{0,s}(\tau_1), \theta^{LG}_{1,s}(\tau_2), \theta^{LG}_{0,s}(\tau_2)).$$

These are estimated by minimizing the sample analogue of the objective function in Theorem C.1. Specifically,

$$\begin{aligned}
& (\tilde{\theta}^{LG}_{1,s}(\tau_1), \tilde{\theta}^{LG}_{0,s}(\tau_1), \tilde{\theta}^{LG}_{1,s}(\tau_2), \tilde{\theta}^{LG}_{0,s}(\tau_2)) \\
= & \mathop{\arg\min}_{\theta_{1,1},\theta_{0,1},\theta_{1,1},\theta_{0,1}} Q_n(s,\tau_1,\theta_{1,1},\theta_{0,1}) + Q_n(s,\tau_2,\theta_{1,2},\theta_{0,2}) - 2\tilde{Q}_n(s,\tau_1,\tau_2,\theta_{1,1},\theta_{0,1},\theta_{1,1},\theta_{0,1}),
\end{aligned}$$

where $Q_n(s,\tau,\theta_1,\theta_0)$ is defined in (5.9) and

$$\begin{aligned}
& \tilde{Q}_n(s,\tau_1,\tau_2,\theta_{1,1},\theta_{0,1},\theta_{1,2},\theta_{0,2}) \\
= & \frac{1}{n(s)} \sum_{i \in I(s)} \left(\frac{\hat{g}_{1,s}(X_i,\theta_{1,1})}{\hat{f}_1(\hat{q}_1(\tau_1))} + \frac{\hat{\pi}(s)}{1-\hat{\pi}(s)} \frac{\hat{g}_{0,s}(X_i,\theta_{0,1})}{\hat{f}_0(\hat{q}_0(\tau_1))} \right) \left(\frac{\hat{g}_{1,s}(X_i,\theta_{1,2})}{\hat{f}_1(\hat{q}_1(\tau_2))} + \frac{\hat{\pi}(s)}{1-\hat{\pi}(s)} \frac{\hat{g}_{0,s}(X_i,\theta_{0,2})}{\hat{f}_0(\hat{q}_0(\tau_2))} \right)
\end{aligned}$$

$$+\frac{1}{n_1(s)}\sum_{i\in I_1(s)}\left(\frac{\hat{g}_{1,s}(X_i,\theta_{1,1})}{\hat{f}_1(\hat{q}_1(\tau_1))}+\frac{\hat{\pi}(s)}{1-\hat{\pi}(s)}\frac{\hat{g}_{0,s}(X_i,\theta_{0,1})}{\hat{f}_0(\hat{q}_0(\tau_1))}\right)\frac{1\{Y_i\le\hat{q}_1(\tau_2)\}}{\hat{f}_1(\hat{q}_1(\tau_2))}$$

$$+\frac{1}{n_1(s)}\sum_{i\in I_1(s)}\left(\frac{\hat{g}_{1,s}(X_i,\theta_{1,2})}{\hat{f}_1(\hat{q}_1(\tau_2))}+\frac{\hat{\pi}(s)}{1-\hat{\pi}(s)}\frac{\hat{g}_{0,s}(X_i,\theta_{0,2})}{\hat{f}_0(\hat{q}_0(\tau_2))}\right)\frac{1\{Y_i\le\hat{q}_1(\tau_1)\}}{\hat{f}_1(\hat{q}_1(\tau_1))}$$

$$+\frac{\hat{\pi}(s)}{n_0(s)(1-\hat{\pi}(s))}\sum_{i\in I_0(s)}\left(\frac{\hat{g}_{1,s}(X_i,\theta_{1,1})}{\hat{f}_1(\hat{q}_1(\tau_1))}+\frac{\hat{\pi}(s)}{1-\hat{\pi}(s)}\frac{\hat{g}_{0,s}(X_i,\theta_{0,1})}{\hat{f}_0(\hat{q}_0(\tau_1))}\right)\frac{1\{Y_i\le\hat{q}_0(\tau_2)\}}{\hat{f}_0(\hat{q}_0(\tau_2))}$$

$$+\frac{\hat{\pi}(s)}{n_0(s)(1-\hat{\pi}(s))}\sum_{i\in I_0(s)}\left(\frac{\hat{g}_{1,s}(X_i,\theta_{1,2})}{\hat{f}_1(\hat{q}_1(\tau_2))}+\frac{\hat{\pi}(s)}{1-\hat{\pi}(s)}\frac{\hat{g}_{0,s}(X_i,\theta_{0,2})}{\hat{f}_0(\hat{q}_0(\tau_2))}\right)\frac{1\{Y_i\le\hat{q}_0(\tau_1)\}}{\hat{f}_0(\hat{q}_0(\tau_1))}.$$

D When $\Theta_s(\tau)$ Is Not a Singleton

Recall the definition of $\Theta_s(\tau)$ in Assumption 9 for the logistic model. In this section we relax the requirement that $\Theta_s(\tau)\cap\Theta$ is a singleton and propose a way to consistently estimate one point, denoted as $\theta_{a,s}(\tau)$, that belongs to the set of optimizers. We work with a fixed τ. The extension of the results to multiple τ's is straightforward. The extension to a continuum of τ's is left for future research.

We need to modify the definition of $\hat{\Theta}_s(\tau)$. Let ε_n be some deterministic sequence such that $\varepsilon_n\downarrow 0$ and

$$\hat{\Theta}_s^{\varepsilon_n}(\tau)=\{(\theta_1,\theta_0)\in\Theta: Q_n(s,\tau,\theta_1,\theta_0)\le\inf_{(\theta_1,\theta_0)\in\Theta}Q_n(s,\tau,\theta_1,\theta_0)+\varepsilon_n\}.$$

Define

$$(\hat{\theta}^*_{1,s}(\tau),\hat{\theta}^*_{0,s}(\tau))=\operatorname*{arg\,min}_{(\theta_1,\theta_0)\in\hat{\Theta}_s^{\varepsilon_n}(\tau)}(||\theta_1||_2^2+||\theta_0||_2^2),$$

and

$$(\theta^*_{1,s}(\tau),\theta^*_{0,s}(\tau))=\operatorname*{arg\,min}_{(\theta_1,\theta_0)\in\Theta_s(\tau)\cap\Theta}(||\theta_1||_2^2+||\theta_0||_2^2).$$

Assumption D.1. *Suppose $(\theta_{1,s}(\tau),\theta_{0,s}(\tau))$ is uniquely defined and $|\hat{f}_a(\hat{q}_a(\tau))-f_a(q_a(\tau))|=o_p(\varepsilon_n)$ for $a=0,1$.*

Proposition D.1. *Suppose Assumptions 1, 2, 7, 8, and D.1 hold. Then, pointwise in τ,*

$$\hat{\theta}^*_{a,s}(\tau)\xrightarrow{p}\theta^*_{a,s}(\tau).$$

E Additional Notation

Throughout the supplement we denote $(\xi_i^s, X_i^s, Y_i^s(1), Y_i^s(0))_{i\in[n]}$ as an i.i.d. sequence with marginal distribution equal to the conditional distribution of $(\xi_i, S_i, Y_i(1), Y_i(0))$ given $S_i = s$. In addition, $\{(\xi_i^s, X_i^s, Y_i^s(1), Y_i^s(0))_{i\in[n]}\}_{s\in\mathcal{S}}$ are independent across s and with $\{A_i, S_i\}_{i\in[n]}$. We further denote $\mathcal{F}$ as a generic class of functions which differs in different contexts. The envelope of $\mathcal{F}$ is denoted as F_i. We say $\mathcal{F}$ is of VC-type with coefficients (α_n, v_n) if

$$\sup_Q N\left(\mathcal{F}, e_Q, \varepsilon||F||_{Q,2}\right) \leq \left(\frac{\alpha_n}{\varepsilon}\right)^{v_n}, \quad \forall \varepsilon \in (0,1],$$

where $N(\cdot)$ denote the covering number, $e_Q(f,g) = ||f-g||_{Q,2}$, and the supremum is taken over all finitely discrete probability measures.

F Proof of Theorem 3.1

We first derive the linear expansion of $\hat{q}_1^{adj}(\tau)$. By Knight's identity (Knight (1998)), we have

$$\begin{aligned}L_n(u,\tau) =& \sum_{i\in[n]}\left[\frac{A_i}{\hat{\pi}(S_i)}\left[\rho_\tau(Y_i - q_1(\tau) - u/\sqrt{n}) - \rho_\tau(Y_i - q_1(\tau))\right] + \frac{(A_i - \hat{\pi}(S_i))}{\hat{\pi}(S_i)\sqrt{n}}\widehat{m}_1(\tau, S_i, X_i)u\right] \\ \equiv& - L_{1,n}(\tau)u + L_{2,n}(u,\tau),\end{aligned}$$

where

$$L_{1,n}(\tau) = \frac{1}{\sqrt{n}}\sum_{i\in[n]}\left[\frac{A_i}{\hat{\pi}(S_i)}(\tau - 1\{Y_i \leq q_1(\tau)\}) - \frac{(A_i - \hat{\pi}(S_i))}{\hat{\pi}(S_i)}\widehat{m}_1(\tau, S_i, X_i)\right]$$

and

$$L_{2,n}(\tau) = \sum_{i\in[n]}\frac{A_i}{\hat{\pi}(S_i)}\int_0^{\frac{u}{\sqrt{n}}}\left(1\{Y_i \leq q_1(\tau)+v\} - 1\{Y_i \leq q_1(\tau)\}\right)dv.$$

By change of variables, we have

$$\sqrt{n}(\hat{q}_1^{adj}(\tau) - q_1(\tau)) = \arg\min_u L_n(u,\tau).$$

Note that $L_{2,n}(\tau)$ is exactly the same as that considered in the proof of Theorem 3.2 in Zhang and Zheng (2020) and by their result we have

$$\sup_{\tau\in\Upsilon}\left|L_{2,n}(\tau) - \frac{f_1(q_1(\tau))u^2}{2}\right| = o_p(1).$$

Next, consider $L_{1,n}(\tau)$. Denote $m_1(\tau,s) = \mathbb{E}(m_1(\tau, S_i, X_i)|S_i = s)$, $\eta_{i,1}(s,\tau) = \tau - 1\{Y_i \leq q_1(\tau)\} -$

$m_1(\tau, s)$, and

$$
\begin{aligned}
L_{1,n}(\tau) =& \frac{1}{\sqrt{n}} \sum_{i \in [n]} \left[\frac{A_i}{\hat{\pi}(S_i)} (\tau - 1\{Y_i \leq q_1(\tau)\}) \right] - \frac{1}{\sqrt{n}} \sum_{i \in [n]} \left[\frac{(A_i - \hat{\pi}(S_i))}{\hat{\pi}(S_i)} \widehat{m}_1(\tau, S_i, X_i) \right] \\
\equiv& L_{1,1,n}(\tau) - L_{1,2,n}(\tau).
\end{aligned}
$$

First, note that $\hat{\pi}(s) - \pi(s) = \frac{D_n(s)}{n(s)}$. Therefore,

$$
\begin{aligned}
& L_{1,1,n}(\tau) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{s \in \mathcal{S}} \frac{A_i}{\pi(s)} 1\{S_i = s\}(\tau - 1\{Y_i(1) \leq q_1(\tau)\}) \\
& - \sum_{i=1}^{n} \sum_{s \in \mathcal{S}} \frac{A_i 1\{S_i = s\}(\hat{\pi}(s) - \pi(s))}{\sqrt{n}\hat{\pi}(s)\pi(s)} (\tau - 1\{Y_i(1) \leq q_1(\tau)\}) \\
& = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{s \in \mathcal{S}} \frac{A_i}{\pi(s)} 1\{S_i = s\}(\tau - 1\{Y_i(1) \leq q_1(\tau)\}) \\
& - \sum_{i=1}^{n} \sum_{s \in \mathcal{S}} \frac{A_i 1\{S_i = s\} D_n(s)}{n(s)\sqrt{n}\hat{\pi}(s)\pi(s)} \eta_{i,1}(s, \tau) - \sum_{s \in \mathcal{S}} \frac{D_n(s) m_1(\tau, s)}{n(s)\sqrt{n}\hat{\pi}(s)\pi(s)} D_n(s) - \sum_{s \in \mathcal{S}} \frac{D_n(s) m_1(\tau, s)}{\sqrt{n}\hat{\pi}(s)} \\
& = \sum_{s \in \mathcal{S}} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{A_i 1\{S_i = s\}}{\pi(s)} \eta_{i,1}(s, \tau) + \sum_{s \in \mathcal{S}} \frac{D_n(s)}{\sqrt{n}\pi(s)} m_1(\tau, s) + \sum_{i=1}^{n} \frac{m_1(\tau, S_i)}{\sqrt{n}} \\
& - \sum_{i=1}^{n} \sum_{s \in \mathcal{S}} \frac{A_i 1\{S_i = s\} D_n(s)}{n(s)\sqrt{n}\hat{\pi}(s)\pi(s)} \eta_{i,1}(s, \tau) - \sum_{s \in \mathcal{S}} \frac{D_n(s) m_1(s, \tau)}{n(s)\sqrt{n}\hat{\pi}(s)\pi(s)} D_n(s) - \sum_{s \in \mathcal{S}} \frac{D_n(s) m_1(\tau, s)}{\sqrt{n}\hat{\pi}(s)} \\
& = \sum_{s \in \mathcal{S}} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{A_i 1\{S_i = s\}}{\pi(s)} \eta_{i,1}(s, \tau) + \sum_{i=1}^{n} \frac{m_1(\tau, S_i)}{\sqrt{n}} + R_{1,1}(\tau), \qquad \text{(F.1)}
\end{aligned}
$$

where

$$
\begin{aligned}
R_{1,1}(\tau) =& - \sum_{i=1}^{n} \sum_{s \in \mathcal{S}} \frac{A_i 1\{S_i = s\} D_n(s)}{n(s)\sqrt{n}\hat{\pi}(s)\pi(s)} \eta_{i,1}(s, \tau) - \sum_{s \in \mathcal{S}} \frac{D_n(s) m_1(\tau, s)}{n(s)\sqrt{n}\hat{\pi}(s)\pi(s)} D_n(s) \\
& + \sum_{s \in \mathcal{S}} \frac{D_n(s) m_1(\tau, s)}{\sqrt{n}} \left(\frac{1}{\pi(s)} - \frac{1}{\hat{\pi}(s)} \right) \\
=& - \sum_{i=1}^{n} \sum_{s \in \mathcal{S}} \frac{A_i 1\{S_i = s\} D_n(s)}{n(s)\sqrt{n}\hat{\pi}(s)\pi(s)} \eta_{i,1}(s, \tau).
\end{aligned}
$$

In addition, note that

$$
\{\tau - 1\{Y_i(1) \leq q_1(\tau)\} - m_1(\tau, S_i) : \tau \in \Upsilon\}
$$

is of the VC-type with fixed coefficients (α, v) and bounded envelope, and $\mathbb{E}(\tau - 1\{Y_i(1) \leq q_1(\tau)\} -$

$m_1(\tau, S_i)|S_i = s) = 0$. Therefore, Lemma P.2 implies

$$\sup_{\tau\in\Upsilon, s\in\mathcal{S}} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} A_i 1\{S_i = s\} \eta_{i,1}(s,\tau) \right| = O_p(1).$$

By Assumption 1 we have $\max_{s\in\mathcal{S}} |D_n(s)/n(s)| = o_p(1)$, $\max_{s\in\mathcal{S}} |\hat{\pi}(s) - \pi(s)| = o_p(1)$, and $\min_{s\in\mathcal{S}} \pi(s) > c > 0$, which imply $\sup_{\tau\in\Upsilon} |R_{1,1}(\tau)| = o_p(1)$.

Next, denote $\overline{m}_1(\tau, s) = \mathbb{E}(\overline{m}_1(q_1(\tau), X_i, S_i)|S_i = s)$. Then

$$\begin{aligned}
L_{1,2,n} &= \frac{1}{\sqrt{n}} \sum_{s\in\mathcal{S}} \sum_{i\in[n]} \frac{A_i}{\hat{\pi}(s)} \overline{m}_1(\tau, s, X_i) 1\{S_i = s\} - \frac{1}{\sqrt{n}} \sum_{i\in[n]} \overline{m}_1(\tau, S_i, X_i) \\
&+ \frac{1}{\sqrt{n}} \sum_{s\in\mathcal{S}} \frac{1}{\hat{\pi}(s)} \sum_{i\in[n]} (A_i - \hat{\pi}(s)) \left(\widehat{m}_1(\tau, s, X_i) - \overline{m}_1(\tau, s, X_i)\right) 1\{S_i = s\} \\
&= \frac{1}{\sqrt{n}} \sum_{s\in\mathcal{S}} \sum_{i\in[n]} \frac{A_i}{\hat{\pi}(s)} (\overline{m}_1(\tau, s, X_i) - \overline{m}_1(\tau, s)) 1\{S_i = s\} \\
&- \frac{1}{\sqrt{n}} \sum_{i\in[n]} \left(\overline{m}_1(\tau, S_i, X_i) - \overline{m}_1(\tau, S_i)\right) \\
&+ \frac{1}{\sqrt{n}} \sum_{s\in\mathcal{S}} \frac{1}{\hat{\pi}(s)} \sum_{i\in[n]} (A_i - \hat{\pi}(s)) \left(\widehat{m}_1(\tau, s, X_i) - \overline{m}_1(\tau, s, X_i)\right) 1\{S_i = s\} \\
&= \frac{1}{\sqrt{n}} \sum_{s\in\mathcal{S}} \sum_{i\in[n]} \frac{A_i}{\pi(s)} (\overline{m}_1(\tau, s, X_i) - \overline{m}_1(\tau, s)) 1\{S_i = s\} \\
&- \frac{1}{\sqrt{n}} \sum_{i\in[n]} \left(\overline{m}_1(\tau, S_i, X_i) - \overline{m}_1(\tau, S_i)\right) \\
&+ \frac{1}{\sqrt{n}} \sum_{s\in\mathcal{S}} \left(\frac{\pi(s) - \hat{\pi}(s)}{\hat{\pi}(s)\pi(s)}\right) \left(\sum_{i\in[n]} A_i (\overline{m}_1(\tau, s, X_i) - \overline{m}_1(\tau, s)) 1\{S_i = s\}\right) \\
&+ \frac{1}{\sqrt{n}} \sum_{s\in\mathcal{S}} \frac{1}{\hat{\pi}(s)} \sum_{i\in[n]} (A_i - \hat{\pi}(s)) \left(\widehat{m}_1(\tau, s, X_i) - \overline{m}_1(\tau, s, X_i)\right) 1\{S_i = s\} \\
&\equiv \frac{1}{\sqrt{n}} \sum_{s\in\mathcal{S}} \sum_{i\in[n]} \frac{A_i}{\pi(s)} (\overline{m}_1(\tau, s, X_i) - \overline{m}_1(\tau, s)) 1\{S_i = s\} \\
&- \frac{1}{\sqrt{n}} \sum_{i\in[n]} \left(\overline{m}_1(\tau, S_i, X_i) - \overline{m}_1(\tau, S_i)\right) + R_2(\tau),
\end{aligned} \tag{F.2}$$

where the second equality holds because

$$\sum_{s\in\mathcal{S}} \sum_{i\in[n]} \frac{A_i}{\hat{\pi}(s)} \overline{m}_1(\tau, s) 1\{S_i = s\} = \sum_{s\in\mathcal{S}} n(s) \overline{m}_1(\tau, s) = \sum_{i\in[n]} \overline{m}_1(\tau, S_i).$$

For the first term of $R_2(\tau)$, we have

$$\begin{aligned}
&\sup_{\tau\in\Upsilon}\left|\frac{1}{\sqrt{n}}\sum_{s\in\mathcal{S}}\left(\frac{\pi(s)-\hat{\pi}(s)}{\hat{\pi}(s)\pi(s)}\right)\left(\sum_{i\in[n]}A_i(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s))1\{S_i=s\}\right)\right| \\
&\leq \sum_{s\in\mathcal{S}}\left|\frac{D_n(s)}{n_1(s)\pi(s)}\right|\sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{1}{\sqrt{n}}\sum_{i\in[n]}A_i1\{S_i=s\}(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s))\right|.
\end{aligned}$$

Assumption 3 implies

$$\mathcal{F}=\{\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s):\tau\in\Upsilon\}$$

is of the VC-type with fixed coefficients (α,v) and an envelope F_i such that $\mathbb{E}(|F_i|^q|S_i=s)<\infty$ for $q>2$. Therefore,

$$\sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{1}{\sqrt{n}}\sum_{i\in[n]}A_i1\{S_i=s\}(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s))\right|=O_p(n^{-1/2}).$$

It is also assumed that $D_n(s)/n(s)=o_p(1)$ and $n(s)/n_1(s)\xrightarrow{p}1/\pi(s)<\infty$. Therefore, we have

$$\sup_{\tau\in\Upsilon}\left|\frac{1}{\sqrt{n}}\sum_{s\in\mathcal{S}}\left(\frac{\pi(s)-\hat{\pi}(s)}{\hat{\pi}(s)\pi(s)}\right)\left(\sum_{i\in[n]}A_i(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s))1\{S_i=s\}\right)\right|=o_p(1).$$

Recall $\overline{\Delta}_1(\tau,s,X_i)=\widehat{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s,X_i)$. Then, for the second term of $R_2(\tau)$, we have

$$\begin{aligned}
&\left|\frac{1}{\sqrt{n}}\sum_{s\in\mathcal{S}}\frac{1}{\hat{\pi}(s)}\sum_{i\in[n]}(A_i-\hat{\pi}(s))\overline{\Delta}_1(\tau,s,X_i)1\{S_i=s\}\right| \\
=&\frac{1}{\sqrt{n}}\sum_{s\in\mathcal{S}}n_0(s)\sup_{\tau\in\Upsilon}\left|\frac{\sum_{i\in I_1(s)}\overline{\Delta}_1(\tau,s,X_i)}{n_1(s)}-\frac{\sum_{i\in I_0(s)}\overline{\Delta}_1(\tau,s,X_i)}{n_0(s)}\right|=o_p(1),
\end{aligned}$$

where the last equality holds by Assumption 3(i). Therefore, we have

$$\sup_{\tau\in\Upsilon}|R_{1,2}(\tau)|=o_p(1).$$

Combining (F.1) and (F.2), we have

$$\begin{aligned}
L_{1,n}(\tau)=&\sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}A_i1\{S_i=s\}\left[\frac{\eta_{i,1}(s,\tau)}{\pi(s)}+\left(1-\frac{1}{\pi(s)}\right)(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s))\right] \\
&+\sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(1-A_i)1\{S_i=s\}\left(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s)\right)
\end{aligned}$$

$$+\sum_{i=1}^n \frac{m_1(\tau,S_i)}{\sqrt{n}} + R_{1,1}(\tau) - R_{1,2}(\tau).$$

Note by Assumption 3 that the classes of functions

$$\left\{\frac{\eta_{i,1}(s,\tau)}{\pi(s)} + \left(1-\frac{1}{\pi(s)}\right)(\overline{m}_1(\tau,s,X_i) - \overline{m}_1(\tau,s)) : \tau \in \Upsilon\right\}$$

and

$$\{\overline{m}_1(\tau,s,X_i) - \overline{m}_1(\tau,s) : \tau \in \Upsilon\}$$

are of the VC-type with fixed coefficients and envelopes belonging to $L_{\mathbb{P},q}$. In addition,

$$\mathbb{E}\left[\frac{\eta_{i,1}(s,\tau)}{\pi(s)} + \left(1-\frac{1}{\pi(s)}\right)(\overline{m}_1(\tau,s,X_i) - \overline{m}_1(\tau,s))|S_i = s\right] = 0$$

and

$$\mathbb{E}(\overline{m}_1(\tau,s,X_i) - \overline{m}_1(\tau,s)|S_i = s) = 0.$$

Therefore, Lemma P.2 implies,

$$\sup_{\tau\in\Upsilon}\left|\sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i=1}^n A_i 1\{S_i = s\}\left[\frac{\eta_{i,1}(s,\tau)}{\pi(s)} + \left(1-\frac{1}{\pi(s)}\right)(\overline{m}_1(\tau,s,X_i) - \overline{m}_1(\tau,s))\right]\right| = O_p(1)$$

and

$$\sup_{\tau\in\Upsilon}\left|\sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i=1}^n (1-A_i) 1\{S_i = s\}(\overline{m}_1(\tau,s,X_i) - \overline{m}_1(\tau,s))\right| = O_p(1).$$

This implies $\sup_{\tau\in\Upsilon}|L_{1,n}(\tau)| = O_p(1)$. Then by Kato (2009, Theorem 2), we have

$$\begin{aligned}
&\sqrt{n}(\hat{q}_1^{adj}(\tau) - q_1(\tau))\\
&= \frac{1}{f_1(q_1(\tau))}\left\{\sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i=1}^n A_i 1\{S_i = s\}\left[\frac{\eta_{i,1}(s,\tau)}{\pi(s)} + \left(1-\frac{1}{\pi(s)}\right)(\overline{m}_1(\tau,s,X_i) - \overline{m}_1(\tau,s))\right]\right.\\
&\left.+\sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i=1}^n (1-A_i) 1\{S_i = s\}(\overline{m}_1(\tau,s,X_i) - \overline{m}_1(\tau,s)) + \sum_{i=1}^n \frac{m_1(\tau,S_i)}{\sqrt{n}}\right\} + R_{q,1}(\tau),
\end{aligned}$$

where $\sup_{\tau\in\Upsilon}|R_{q,1}(\tau)| = o_p(1)$. Similarly, we have

$$\begin{aligned}
&\sqrt{n}(\hat{q}_0^{adj}(\tau) - q_0(\tau))\\
=&\frac{1}{f_0(q_0(\tau))}\left\{\sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i=1}^n (1-A_i) 1\{S_i = s\}\left[\frac{\eta_{i,0}(s,\tau)}{1-\pi(s)} + \left(1-\frac{1}{1-\pi(s)}\right)(\overline{m}_0(\tau,s,X_i) - \overline{m}_0(\tau,s))\right]\right.
\end{aligned}$$

$$+\sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}A_i1\{S_i=s\}\left(\overline{m}_0(\tau,s,X_i)-\overline{m}_0(\tau,s)\right)+\sum_{i=1}^{n}\frac{m_0(\tau,S_i)}{\sqrt{n}}\Bigg\}+R_{q,0}(\tau),$$

where $\eta_{i,0}(s,\tau)=\tau-1\{Y_i(0)\leq q_0(\tau)\}-m_0(\tau,s)$ and $\sup_{\tau\in\Upsilon}|R_{q,0}(\tau)|=o_p(1)$. Taking the difference of the above two displays gives

$$\begin{aligned}
&\sqrt{n}(\hat{q}^{adj}(\tau)-q(\tau))\\
&=\sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i\in[n]}A_i1\{S_i=s\}\left[\frac{\eta_{i,1}(s,\tau)-(1-\pi(s))\left(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s)\right)}{\pi(s)f_1(q_1(\tau))}-\frac{\left(\overline{m}_0(\tau,s,X_i)-\overline{m}_0(\tau,s)\right)}{f_0(q_0(\tau))}\right]\\
&-\sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i\in[n]}(1-A_i)1\{S_i=s\}\left[\frac{\eta_{i,0}(s,\tau)-\pi(s)\left(\overline{m}_0(\tau,s,X_i)-\overline{m}_0(\tau,s)\right)}{(1-\pi(s))f_0(q_0(\tau))}-\frac{\left(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s)\right)}{f_1(q_1(\tau))}\right]\\
&+\frac{1}{\sqrt{n}}\sum_{i\in[n]}\left(\frac{m_1(\tau,S_i)}{f_1(q_1(\tau))}-\frac{m_0(\tau,S_i)}{f_0(q_0(\tau))}\right)+R_q(\tau)\\
&\equiv\sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i\in[n]}A_i1\{S_i=s\}\phi_1(\tau,s,Y_i(1),X_i)-\sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i\in[n]}(1-A_i)1\{S_i=s\}\phi_0(\tau,s,Y_i(0),X_i)\\
&+\frac{1}{\sqrt{n}}\sum_{i\in[n]}\phi_s(\tau,S_i)+R_q(\tau),
\end{aligned}$$

where $\sup_{\tau\in\Upsilon}|R_q(\tau)|=o_p(1)$. Lemma P.3 shows that, uniformly over $\tau\in\Upsilon$,

$$\sqrt{n}(\hat{q}^{adj}(\tau)-q(\tau))\rightsquigarrow\mathcal{B}(\tau),$$

where $\mathcal{B}(\tau)$ is a Gaussian process with covariance kernel

$$\begin{aligned}
\Sigma(\tau,\tau')=&\,\mathbb{E}\pi(S_i)\phi_1(\tau,S_i,Y_i(1),X_i)\phi_1(\tau',S_i,Y_i(1),X_i)\\
&+\mathbb{E}(1-\pi(S_i))\phi_0(\tau,S_i,Y_i(0),X_i)\phi_0(\tau',S_i,Y_i(0),X_i)\\
&+\mathbb{E}\phi_s(\tau,S_i)\phi_s(\tau',S_i).
\end{aligned}$$

For the second result in Theorem 3.1, we denote

$$\delta_a(\tau,S_i,X_i)=m_a(\tau,S_i,X_i)-m_a(\tau,S_i)\quad\text{and}\quad\overline{\delta}_a(\tau,S_i,X_i)=(\overline{m}_a(\tau,S_i,X_i)-\overline{m}_a(\tau,S_i)),\ a=0,1. \tag{F.3}$$

Then

$$\begin{aligned}
&\mathbb{E}\pi(S_i)\phi_1(\tau,S_i,Y_i(1),X_i)\phi_1(\tau',S_i,Y_i(1),X_i)\\
&=\mathbb{E}\frac{1}{\pi(S_i)}\left[\frac{(\tau-1\{Y_i(1)\leq q_1(\tau)\}-m_1(\tau,S_i,X_i))}{f_1(q_1(\tau))}\right]\left[\frac{(\tau'-1\{Y_i(1)\leq q_1(\tau')\}-m_1(\tau',S_i,X_i))}{f_1(q_1(\tau'))}\right]
\end{aligned}$$

$$
\begin{aligned}
&+ \mathbb{E}\pi(S_i) \left[\frac{\delta_1(\tau, S_i, X_i) - \overline{\delta}_1(\tau, S_i, X_i)}{\pi(S_i) f_1(q_1(\tau))} + \left(\frac{\overline{\delta}_1(\tau, S_i, X_i)}{f_1(q_1(\tau))} - \frac{\overline{\delta}_0(\tau, S_i, X_i)}{f_0(q_0(\tau))} \right) \right] \\
&\times \left[\frac{\delta_1(\tau', S_i, X_i) - \overline{\delta}_1(\tau', S_i, X_i)}{\pi(S_i) f_1(q_1(\tau'))} + \left(\frac{\overline{\delta}_1(\tau', S_i, X_i)}{f_1(q_1(\tau'))} - \frac{\overline{\delta}_0(\tau', S_i, X_i)}{f_0(q_0(\tau'))} \right) \right],
\end{aligned}
$$

$$
\begin{aligned}
&\mathbb{E}(1-\pi(S_i))\phi_0(\tau, S_i, Y_i(1), X_i)\phi_0(\tau', S_i, Y_i(1), X_i) \\
&= \mathbb{E}\frac{1}{1-\pi(S_i)} \left[\frac{(\tau - 1\{Y_i(0) \le q_0(\tau)\} - m_0(\tau, S_i, X_i))}{f_1(q_1(\tau))} \right] \left[\frac{(\tau' - 1\{Y_i(0) \le q_0(\tau')\} - m_0(\tau', S_i, X_i))}{f_1(q_1(\tau'))} \right] \\
&+ \mathbb{E}(1-\pi(S_i)) \left[\frac{\delta_0(\tau, S_i, X_i) - \overline{\delta}_0(\tau, S_i, X_i)}{(1-\pi(S_i)) f_0(q_0(\tau))} - \left(\frac{\overline{\delta}_1(\tau, S_i, X_i)}{f_1(q_1(\tau))} - \frac{\overline{\delta}_0(\tau, S_i, X_i)}{f_0(q_0(\tau))} \right) \right] \\
&\times \left[\frac{\delta_0(\tau', S_i, X_i) - \overline{\delta}_0(\tau', S_i, X_i)}{(1-\pi(S_i)) f_0(q_0(\tau'))} - \left(\frac{\overline{\delta}_1(\tau', S_i, X_i)}{f_1(q_1(\tau'))} - \frac{\overline{\delta}_0(\tau', S_i, X_i)}{f_0(q_0(\tau'))} \right) \right],
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{E}\phi_s(\tau, S_i)\phi_s(\tau', S_i) &= \mathbb{E}\left(\frac{m_1(\tau, S_i)}{f_1(q_1(\tau))} - \frac{m_0(\tau, S_i)}{f_0(q_0(\tau))} \right) \left(\frac{m_1(\tau', S_i)}{f_1(q_1(\tau'))} - \frac{m_0(\tau', S_i)}{f_0(q_0(\tau'))} \right) \\
&= \mathbb{E}\left(\frac{m_1(\tau, S_i, X_i)}{f_1(q_1(\tau))} - \frac{m_0(\tau, S_i, X_i)}{f_0(q_0(\tau))} \right) \left(\frac{m_1(\tau', S_i, X_i)}{f_1(q_1(\tau'))} - \frac{m_0(\tau', S_i, X_i)}{f_0(q_0(\tau'))} \right) \\
&+ \mathbb{E}\left(\frac{\delta_1(\tau, S_i, X_i)}{f_1(q_1(\tau))} - \frac{\delta_0(\tau, S_i, X_i)}{f_0(q_0(\tau))} \right) \left(\frac{\delta_1(\tau', S_i, X_i)}{f_1(q_1(\tau'))} - \frac{\delta_0(\tau', S_i, X_i)}{f_0(q_0(\tau'))} \right).
\end{aligned}
$$

Let

$$
\begin{aligned}
&\Sigma^*(\tau, \tau') \\
&= \mathbb{E}\frac{1}{\pi(S_i)} \left[\frac{(\tau - 1\{Y_i(1) \le q_1(\tau)\} - m_1(\tau, S_i, X_i))}{f_1(q_1(\tau))} \right] \left[\frac{(\tau' - 1\{Y_i(1) \le q_1(\tau')\} - m_1(\tau', S_i, X_i))}{f_1(q_1(\tau'))} \right] \\
&+ \mathbb{E}\frac{1}{1-\pi(S_i)} \left[\frac{(\tau - 1\{Y_i(0) \le q_0(\tau)\} - m_0(\tau, S_i, X_i))}{f_1(q_1(\tau))} \right] \left[\frac{(\tau' - 1\{Y_i(0) \le q_0(\tau')\} - m_0(\tau', S_i, X_i))}{f_1(q_1(\tau'))} \right] \\
&+ \mathbb{E}\left(\frac{m_1(\tau, S_i, X_i)}{f_1(q_1(\tau))} - \frac{m_0(\tau, S_i, X_i)}{f_0(q_0(\tau))} \right) \left(\frac{m_1(\tau', S_i, X_i)}{f_1(q_1(\tau'))} - \frac{m_0(\tau', S_i, X_i)}{f_0(q_0(\tau'))} \right),
\end{aligned}
$$

which does not rely on the working models. Then,

$$
\begin{aligned}
&\Sigma(\tau, \tau') - \Sigma^*(\tau, \tau') \\
&= \mathbb{E}\pi(S_i) \left[\frac{\delta_1(\tau, S_i, X_i) - \overline{\delta}_1(\tau, S_i, X_i)}{\pi(S_i) f_1(q_1(\tau))} + \left(\frac{\overline{\delta}_1(\tau, S_i, X_i)}{f_1(q_1(\tau))} - \frac{\overline{\delta}_0(\tau, S_i, X_i)}{f_0(q_0(\tau))} \right) \right] \\
&\times \left[\frac{\delta_1(\tau', S_i, X_i) - \overline{\delta}_1(\tau', S_i, X_i)}{\pi(S_i) f_1(q_1(\tau'))} + \left(\frac{\overline{\delta}_1(\tau', S_i, X_i)}{f_1(q_1(\tau'))} - \frac{\overline{\delta}_0(\tau', S_i, X_i)}{f_0(q_0(\tau'))} \right) \right] \\
&+ \mathbb{E}(1-\pi(S_i)) \left[\frac{\delta_0(\tau, S_i, X_i) - \overline{\delta}_0(\tau, S_i, X_i)}{(1-\pi(S_i)) f_0(q_0(\tau))} - \left(\frac{\overline{\delta}_1(\tau, S_i, X_i)}{f_1(q_1(\tau))} - \frac{\overline{\delta}_0(\tau, S_i, X_i)}{f_0(q_0(\tau))} \right) \right]
\end{aligned}
$$

$$
\begin{aligned}
&\times \left[\frac{\delta_0(\tau', S_i, X_i) - \overline{\delta}_0(\tau', S_i, X_i)}{(1-\pi(S_i))f_0(q_0(\tau'))} - \left(\frac{\overline{\delta}_1(\tau', S_i, X_i)}{f_1(q_1(\tau'))} - \frac{\overline{\delta}_0(\tau', S_i, X_i)}{f_0(q_0(\tau'))}\right)\right] \\
&- \mathbb{E}\left(\frac{\delta_1(\tau, S_i, X_i)}{f_1(q_1(\tau))} - \frac{\delta_0(\tau, S_i, X_i)}{f_0(q_0(\tau))}\right)\left(\frac{\delta_1(\tau', S_i, X_i)}{f_1(q_1(\tau'))} - \frac{\delta_0(\tau', S_i, X_i)}{f_0(q_0(\tau'))}\right) \\
&= \mathbb{E}\left[\sqrt{\frac{1-\pi(S_i)}{\pi(S_i)}}\frac{\delta_1(\tau, S_i, X_i) - \overline{\delta}_1(\tau, S_i, X_i)}{f_1(q_1(\tau))} + \sqrt{\frac{\pi(S_i)}{1-\pi(S_i)}}\frac{\delta_0(\tau, S_i, X_i) - \overline{\delta}_0(\tau, S_i, X_i)}{f_0(q_0(\tau))}\right] \\
&\times \left[\sqrt{\frac{1-\pi(S_i)}{\pi(S_i)}}\frac{\delta_1(\tau', S_i, X_i) - \overline{\delta}_1(\tau', S_i, X_i)}{f_1(q_1(\tau'))} + \sqrt{\frac{\pi(S_i)}{1-\pi(S_i)}}\frac{\delta_0(\tau', S_i, X_i) - \overline{\delta}_0(\tau', S_i, X_i)}{f_0(q_0(\tau'))}\right] \\
&\equiv \mathbb{E}a_i(\tau)a_i(\tau'),
\end{aligned}
$$

where

$$
a_i(\tau) = \sqrt{\frac{1-\pi(S_i)}{\pi(S_i)}}\frac{\delta_1(\tau, S_i, X_i) - \overline{\delta}_1(\tau, S_i, X_i)}{f_1(q_1(\tau))} + \sqrt{\frac{\pi(S_i)}{1-\pi(S_i)}}\frac{\delta_0(\tau, S_i, X_i) - \overline{\delta}_0(\tau, S_i, X_i)}{f_0(q_0(\tau))}.
$$

Further, denote $\vec{a}_i = (a_i(\tau_1), \cdots, a_i(\tau_K))^\top$, the asymptotic variance covariance matrix of $(\hat{q}^{adj}(\tau_1), \cdots, \hat{q}^{adj}(\tau_K))$ as $[\Sigma_{kl}]_{k,l\in[K]}$, and the optimal variance covariance matrix as $[\Sigma^*_{kl}]_{k,l\in[K]}$. We have

$$
[\Sigma_{kl}]_{k,l\in[K]} - [\Sigma^*_{kl}]_{k,l\in[K]} = [\mathbb{E}a_i(\tau_k)a_i(\tau_l)]_{k,l\in[K]} = \mathbb{E}\vec{a}_i\vec{a}_i^\top,
$$

which is positive semidefinite. In addition, $\mathbb{E}\vec{a}_i\vec{a}_i^\top = 0$ if $\overline{m}_a(\tau, s, x) = m_a(\tau, s, x)$ for $a = 0, 1$, $\tau \in \{\tau_1, \cdots, \tau_K\}$, and (s, x) in the joint support of (S_i, X_i). This concludes the proof.

G Proof of Theorem 4.1

We focus on deriving the linear expansion of $\hat{q}^w_1(\tau)$. Let

$$
\begin{aligned}
L^w_n(u, \tau) &= \sum_{i\in[n]} \xi_i \left[\frac{A_i}{\hat{\pi}^w(S_i)}\left[\rho_\tau(Y_i - q_1(\tau) - u/\sqrt{n}) - \rho_\tau(Y_i - q_1(\tau))\right] + \frac{(A_i - \hat{\pi}^w(S_i))}{\hat{\pi}^w(S_i)\sqrt{n}}\widehat{m}_1(\tau, S_i, X_i)u\right] \\
&\equiv -L^w_{1,n}(\tau)u + L^w_{2,n}(u, \tau),
\end{aligned}
$$

where

$$
L^w_{1,n}(\tau) = \frac{1}{\sqrt{n}}\sum_{i\in[n]}\xi_i\left[\frac{A_i}{\hat{\pi}^w(S_i)}(\tau - 1\{Y_i \le q_1(\tau)\}) - \frac{(A_i - \hat{\pi}^w(S_i))}{\hat{\pi}^w(S_i)}\widehat{m}_1(\tau, S_i, X_i)\right],
$$

and

$$
L^w_{2,n}(\tau) = \sum_{i\in[n]}\frac{\xi_i A_i}{\hat{\pi}^w(S_i)}\int_0^{\frac{u}{\sqrt{n}}}\left(1\{Y_i \le q_1(\tau) + v\} - 1\{Y_i \le q_1(\tau)\}\right)dv.
$$

By the change of variables, we have

$$\sqrt{n}(\hat{q}_1^w(\tau) - q_1(\tau)) = \arg\min_u L_n^w(u,\tau).$$

Note that $L_{2,n}^w(\tau)$ is exactly the same as that considered in the proof of Theorem 3.2 in Zhang and Zheng (2020) and by their result we have

$$\sup_{\tau\in\Upsilon}\left|L_{2,n}^w(\tau) - \frac{f_1(q_1(\tau))u^2}{2}\right| = o_p(1).$$

Next consider $L_{1,n}^w(\tau)$. Recall $m_1(\tau,s) = \mathbb{E}(m_1(\tau,S_i,X_i)|S_i = s)$ and $\eta_{i,1}(s,\tau) = \tau - 1\{Y_i \leq q_1(\tau)\} - m_1(\tau,s)$. Denote

$$\begin{aligned}
L_{1,n}^w(\tau) &= \frac{1}{\sqrt{n}}\sum_{i\in[n]}\xi_i\left[\frac{A_i}{\hat{\pi}^w(S_i)}(\tau - 1\{Y_i \leq q_1(\tau)\})\right] - \frac{1}{\sqrt{n}}\sum_{i\in[n]}\xi_i\left[\frac{(A_i - \hat{\pi}^w(S_i))}{\hat{\pi}^w(S_i)}\widehat{m}_1(\tau,S_i,X_i)\right] \\
&\equiv L_{1,1,n}^w(\tau) - L_{1,2,n}^w(\tau).
\end{aligned}$$

First, note that

$$\begin{aligned}
& L_{1,1,n}^w(\tau) = \frac{1}{\sqrt{n}}\sum_{i=1}^n\sum_{s\in\mathcal{S}}\frac{\xi_i A_i}{\pi(s)}1\{S_i = s\}(\tau - 1\{Y_i(1) \leq q_1(\tau)\}) \\
& - \sum_{i=1}^n\sum_{s\in\mathcal{S}}\frac{\xi_i A_i 1\{S_i = s\}(\hat{\pi}^w(s) - \pi(s))}{\sqrt{n}\hat{\pi}^w(s)\pi(s)}(\tau - 1\{Y_i(1) \leq q_1(\tau)\}) \\
& = \frac{1}{\sqrt{n}}\sum_{i=1}^n\sum_{s\in\mathcal{S}}\frac{\xi_i A_i}{\pi(s)}1\{S_i = s\}(\tau - 1\{Y_i(1) \leq q_1(\tau)\}) \\
& - \sum_{i=1}^n\sum_{s\in\mathcal{S}}\frac{\xi_i A_i 1\{S_i = s\}D_n^w(s)}{n^w(s)\sqrt{n}\hat{\pi}^w(s)\pi(s)}\eta_{i,1}(s,\tau) - \sum_{s\in\mathcal{S}}\frac{D_n^w(s)m_1(\tau,s)}{n^w(s)\sqrt{n}\hat{\pi}^w(s)\pi(s)}D_n^w(s) - \sum_{s\in\mathcal{S}}\frac{D_n^w(s)m_1(\tau,s)}{\sqrt{n}\hat{\pi}^w(s)} \\
& = \sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i=1}^n\frac{\xi_i A_i 1\{S_i = s\}}{\pi(s)}\eta_{i,1}(s,\tau) + \sum_{s\in\mathcal{S}}\frac{D_n^w(s)}{\sqrt{n}\pi(s)}m_1(\tau,s) + \sum_{i=1}^n\frac{\xi_i m_1(\tau,S_i)}{\sqrt{n}} \\
& - \sum_{i=1}^n\sum_{s\in\mathcal{S}}\frac{\xi_i A_i 1\{S_i = s\}D_n^w(s)}{n^w(s)\sqrt{n}\hat{\pi}^w(s)\pi(s)}\eta_{i,1}(s,\tau) - \sum_{s\in\mathcal{S}}\frac{D_n^w(s)m_1(s,\tau)}{n^w(s)\sqrt{n}\hat{\pi}^w(s)\pi(s)}D_n^w(s) - \sum_{s\in\mathcal{S}}\frac{D_n^w(s)m_1(\tau,s)}{\sqrt{n}\hat{\pi}^w(s)} \\
& = \sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i=1}^n\frac{\xi_i A_i 1\{S_i = s\}}{\pi(s)}\eta_{i,1}(s,\tau) + \sum_{i=1}^n\frac{\xi_i m_1(\tau,S_i)}{\sqrt{n}} + R_{1,1}^w(\tau), \qquad (G.1)
\end{aligned}$$

where $D_n^w(s) = \sum_{i\in[n]}\xi_i(A_i - \pi(S_i))1\{S_i = s\} = (\pi^w(s) - \pi(s))n^w(s)$,

$$R_{1,1}^w(\tau) = -\sum_{i=1}^n\sum_{s\in\mathcal{S}}\frac{\xi_i A_i 1\{S_i = s\}D_n^w(s)}{n^w(s)\sqrt{n}\hat{\pi}^w(s)\pi(s)}\eta_{i,1}(s,\tau) - \sum_{s\in\mathcal{S}}\frac{D_n^w(s)m_1(\tau,s)}{n^w(s)\sqrt{n}\hat{\pi}^w(s)\pi(s)}D_n^w(s)$$

$$+\sum_{s\in\mathcal{S}}\frac{D_n^w(s)m_1(\tau,s)}{\sqrt{n}}\left(\frac{1}{\pi(s)}-\frac{1}{\hat{\pi}^w(s)}\right)=-\sum_{i=1}^n\sum_{s\in\mathcal{S}}\frac{\xi_iA_i1\{S_i=s\}D_n^w(s)}{n^w(s)\sqrt{n}\hat{\pi}^w(s)\pi(s)}\eta_{i,1}(s,\tau).$$

Note that

$$\{\xi_i(\tau-1\{Y_i(1)\leq q_1(\tau)\}-m_1(\tau,S_i)):\tau\in\Upsilon\}$$

is of the VC-type with fixed coefficients (α,v) and the envelope $F_i=\xi_i$, and

$$\mathbb{E}\left[\xi_i(\tau-1\{Y_i(1)\leq q_1(\tau)\}-m_1(\tau,S_i))|S_i=s\right]=0.$$

We can also let $\sigma_n^2=\mathbb{E}(F_i^2|S_i=s)\leq C<\infty$ for some constant C. Then, Lemma P.2 implies

$$\sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{1}{\sqrt{n}}\sum_{i=1}^nA_i1\{S_i=s\}\xi_i\eta_{i,1}(s,\tau)\right|=O_p(1).$$

In addition, Lemma P.4 implies $\max_{s\in\mathcal{S}}|D_n^w(s)/n^w(s)|=o_p(1)$, which further implies $\max_{s\in\mathcal{S}}|\hat{\pi}^w(s)-\pi(s)|=o_p(1)$. Combining these results, we have

$$\sup_{\tau\in\Upsilon}|R_{1,1}^w(\tau)|=o_p(1).$$

Next, recall $\overline{m}_1(\tau,s)=\mathbb{E}(\overline{m}_1(q_1(\tau),X_i,S_i)|S_i=s)$. Then

$$\begin{aligned}
L_{1,2,n}^w&=\frac{1}{\sqrt{n}}\sum_{s\in\mathcal{S}}\sum_{i\in[n]}\frac{\xi_iA_i}{\hat{\pi}^w(s)}\overline{m}_1(\tau,s,X_i)1\{S_i=s\}-\frac{1}{\sqrt{n}}\sum_{i\in[n]}\xi_i\overline{m}_1(\tau,S_i,X_i)\\
&+\frac{1}{\sqrt{n}}\sum_{s\in\mathcal{S}}\frac{1}{\hat{\pi}^w(s)}\sum_{i\in[n]}\xi_i(A_i-\hat{\pi}^w(s))\left(\widehat{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s,X_i)\right)1\{S_i=s\}\\
&=\frac{1}{\sqrt{n}}\sum_{s\in\mathcal{S}}\sum_{i\in[n]}\frac{\xi_iA_i}{\hat{\pi}^w(s)}(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s))1\{S_i=s\}\\
&-\frac{1}{\sqrt{n}}\sum_{i\in[n]}\xi_i\left(\overline{m}_1(\tau,S_i,X_i)-\overline{m}_1(\tau,S_i)\right)\\
&+\frac{1}{\sqrt{n}}\sum_{s\in\mathcal{S}}\frac{1}{\hat{\pi}^w(s)}\sum_{i\in[n]}\xi_i(A_i-\hat{\pi}^w(s))\left(\widehat{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s,X_i)\right)1\{S_i=s\}\\
&=\frac{1}{\sqrt{n}}\sum_{s\in\mathcal{S}}\sum_{i\in[n]}\frac{\xi_iA_i}{\pi(s)}(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s))1\{S_i=s\}\\
&-\frac{1}{\sqrt{n}}\sum_{i\in[n]}\xi_i\left(\overline{m}_1(\tau,S_i,X_i)-\overline{m}_1(\tau,S_i)\right)\\
&+\frac{1}{\sqrt{n}}\sum_{s\in\mathcal{S}}\left(\frac{\pi(s)-\hat{\pi}^w(s)}{\hat{\pi}^w(s)\pi(s)}\right)\left(\sum_{i\in[n]}\xi_iA_i(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s))1\{S_i=s\}\right)
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{\sqrt{n}}\sum_{s\in\mathcal{S}}\frac{1}{\hat{\pi}^w(s)}\sum_{i\in[n]}\xi_i(A_i-\hat{\pi}^w(s))\left(\widehat{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s,X_i)\right)1\{S_i=s\}\\
&\equiv\frac{1}{\sqrt{n}}\sum_{s\in\mathcal{S}}\sum_{i\in[n]}\frac{\xi_iA_i}{\pi(s)}(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s))1\{S_i=s\}\\
&-\frac{1}{\sqrt{n}}\sum_{i\in[n]}\xi_i\left(\overline{m}_1(\tau,S_i,X_i)-\overline{m}_1(\tau,S_i)\right)+R^w_{1,2}(\tau),
\end{aligned}
\tag{G.2}
$$

where the second equality holds because

$$
\begin{aligned}
\sum_{s\in\mathcal{S}}\sum_{i\in[n]}\xi_iA_i1\{S_i=s\}\frac{\overline{m}_1(\tau,s)}{\hat{\pi}^w(s)}&=\sum_{s\in\mathcal{S}}n_1^w(s)\frac{\overline{m}_1(\tau,s)}{\hat{\pi}^w(s)}=\sum_{s\in\mathcal{S}}n^w(s)\overline{m}_1(\tau,s)\\
&=\sum_{i\in[n]}\sum_{s\in\mathcal{S}}\xi_i1\{S_i=s\}\overline{m}_1(\tau,S_i)=\sum_{i\in[n]}\xi_i\overline{m}_1(\tau,S_i).
\end{aligned}
$$

For the first term in $R^w_2(\tau)$, we have

$$
\begin{aligned}
&\sup_{\tau\in\Upsilon}\left|\frac{1}{\sqrt{n}}\sum_{s\in\mathcal{S}}\left(\frac{\pi(s)-\hat{\pi}^w(s)}{\hat{\pi}^w(s)\pi(s)}\right)\left(\sum_{i\in[n]}A_i\xi_i(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s))1\{S_i=s\}\right)\right|\\
&\leq\sum_{s\in\mathcal{S}}\left|\frac{D_n^w(s)}{n^w(s)\hat{\pi}^w(s)\pi(s)}\right|\sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{1}{\sqrt{n}}\sum_{i\in[n]}\xi_iA_i1\{S_i=s\}(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s))\right|=o_p(1),
\end{aligned}
$$

where the last equality holds due to Lemmas P.2 and P.4, and the fact that $\mathcal{F}=\{\xi(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s)):\tau\in\Upsilon\}$ is of the VC-type with fixed coefficients (α,v) and envelope ξ_iF_i such that $\mathbb{E}((\xi_iF_i)^q|S_i=s)<\infty$ for $q>2$.

For the second term in $R^w_{1,2}(\tau)$, recall $\overline{\Delta}_1(\tau,s,X_i)=\widehat{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s,X_i)$. Then

$$
\begin{aligned}
&\left|\frac{1}{\sqrt{n}}\sum_{s\in\mathcal{S}}\frac{1}{\hat{\pi}^w(s)}\sum_{i\in[n]}\xi_i(A_i-\hat{\pi}^w(s))\overline{\Delta}_1(\tau,s,X_i)1\{S_i=s\}\right|\\
&=\frac{1}{\sqrt{n}}\sum_{s\in\mathcal{S}}n_0^w(s)\sup_{\tau\in\Upsilon}\left|\frac{\sum_{i\in I_1(s)}\xi_i\overline{\Delta}_1(\tau,s,X_i)}{n_1^w(s)}-\frac{\sum_{i\in I_0(s)}\xi_i\overline{\Delta}_1(\tau,s,X_i)}{n_0^w(s)}\right|=o_p(1),
\end{aligned}
$$

where the last equality holds by Assumption 3. Therefore, we have

$$
\sup_{\tau\in\Upsilon}|R^w_{1,2}(\tau)|=o_p(1).
$$

Combining (G.1) and (G.2), we have

$$
L^w_{1,n}(\tau)=\sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i=1}^n\xi_iA_i1\{S_i=s\}\left[\frac{\eta_{i,1}(s,\tau)}{\pi(s)}+\left(1-\frac{1}{\pi(s)}\right)(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s))\right]
$$

$$
\begin{aligned}
&+\sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i(1-A_i)1\{S_i=s\}\left(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s)\right)\\
&+\sum_{i=1}^{n}\xi_i\frac{m_1(\tau,S_i)}{\sqrt{n}}+R^w_{1,1}(\tau)-R^w_{1,2}(\tau),
\end{aligned}
$$

where $\sup_{\tau\in\Upsilon}(|R^w_{1,1}(\tau)|+|R^w_{1,2}(\tau)|)=o_p(1)$. In addition, Assumption 3 implies the classes of functions

$$
\left\{\xi_i\left[\frac{\eta_{i,1}(s,\tau)}{\pi(s)}+\left(1-\frac{1}{\pi(s)}\right)\left(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s)\right)\right]:\tau\in\Upsilon\right\}
$$

and

$$
\{\xi_i[\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s)]:\tau\in\Upsilon\}
$$

are of the VC-type with fixed coefficients and envelopes belonging to $L_{\mathbb{P},q}$. In addition,

$$
\mathbb{E}\left[\xi_i\left(\frac{\eta_{i,1}(s,\tau)}{\pi(s)}+\left(1-\frac{1}{\pi(s)}\right)\left(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s)\right)\right)\Big|S_i=s\right]=0,
$$

and

$$
\mathbb{E}\left[\xi_i(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s))|S_i=s\right]=0.
$$

Therefore, Lemma P.2 implies,

$$
\sup_{\tau\in\Upsilon}\left|\sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}A_i1\{S_i=s\}\left[\frac{\eta_{i,1}(s,\tau)}{\pi(s)}+\left(1-\frac{1}{\pi(s)}\right)\left(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s)\right)\right]\right|=O_p(1),
$$

and

$$
\sup_{\tau\in\Upsilon}\left|\sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(1-A_i)1\{S_i=s\}\left(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s)\right)\right|=O_p(1).
$$

This implies $\sup_{\tau\in\Upsilon}|L^w_{1,n}(\tau)|=O_p(1)$. Then by Kato (2009, Theorem 2) we have

$$
\begin{aligned}
&\sqrt{n}(\hat{q}^w_1(\tau)-q_1(\tau))\\
&=\frac{1}{f_1(q_1(\tau))}\left\{\sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_iA_i1\{S_i=s\}\left[\frac{\eta_{i,1}(s,\tau)}{\pi(s)}+\left(1-\frac{1}{\pi(s)}\right)\left(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s)\right)\right]\right.\\
&\left.+\sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i(1-A_i)1\{S_i=s\}\left(\overline{m}_1(\tau,s,X_i)-\overline{m}_1(\tau,s)\right)+\sum_{i=1}^{n}\xi_i\frac{m_1(\tau,S_i)}{\sqrt{n}}\right\}+R^w_{q,1}(\tau),
\end{aligned}
$$

where $\sup_{\tau\in\Upsilon}|R^w_{q,1}(\tau)| = o_p(1)$. Similarly, we have

$$
\begin{aligned}
&\sqrt{n}(\hat{q}^w_0(\tau) - q_0(\tau)) \\
&= \frac{1}{f_0(q_0(\tau))}\Bigg\{\sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i=1}^n \xi_i(1-A_i)1\{S_i=s\}\left[\frac{\eta_{i,0}(s,\tau)}{1-\pi(s)} + \left(1-\frac{1}{1-\pi(s)}\right)(\overline{m}_0(\tau,s,X_i) - \overline{m}_0(\tau,s))\right] \\
&+ \sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i=1}^n \xi_i A_i 1\{S_i=s\}(\overline{m}_0(\tau,s,X_i) - \overline{m}_0(\tau,s)) + \sum_{i=1}^n \xi_i\frac{m_0(\tau,S_i)}{\sqrt{n}}\Bigg\} + R_{q,0}(\tau),
\end{aligned}
$$

where $\sup_{\tau\in\Upsilon}|R^w_{q,0}(\tau)| = o_p(1)$. Taking the difference of the above two displays we obtain

$$
\begin{aligned}
&\sqrt{n}(\hat{q}^w(\tau) - q(\tau)) \\
&= \sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i\in[n]}\xi_i A_i 1\{S_i=s\} \\
&\times\left[\frac{\eta_{i,1}(s,\tau) - (1-\pi(s))(\overline{m}_1(\tau,s,X_i) - \overline{m}_1(\tau,s))}{\pi(s)f_1(q_1(\tau))} - \frac{(\overline{m}_0(\tau,s,X_i) - \overline{m}_0(\tau,s))}{f_0(q_0(\tau))}\right] \\
&- \sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i\in[n]}\xi_i(1-A_i)1\{S_i=s\} \\
&\times\left[\frac{\eta_{i,0}(s,\tau) - \pi(s)(\overline{m}_0(\tau,s,X_i) - \overline{m}_0(\tau,s))}{(1-\pi(s))f_0(q_0(\tau))} - \frac{(\overline{m}_1(\tau,s,X_i) - \overline{m}_1(\tau,s))}{f_1(q_1(\tau))}\right] \\
&+ \frac{1}{\sqrt{n}}\sum_{i\in[n]}\xi_i\left(\frac{m_1(\tau,S_i)}{f_1(q_1(\tau))} - \frac{m_0(\tau,S_i)}{f_0(q_0(\tau))}\right) + R^w_q(\tau) \\
&= \sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i\in[n]}\xi_i A_i 1\{S_i=s\}\phi_1(\tau,s,Y_i(1),X_i) - \sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i\in[n]}\xi_i(1-A_i)1\{S_i=s\}\phi_0(\tau,s,Y_i(0),X_i) \\
&+ \frac{1}{\sqrt{n}}\sum_{i\in[n]}\xi_i\phi_s(\tau,S_i) + R^w_q(\tau),
\end{aligned}
$$

where $\sup_{\tau\in\Upsilon}|R^w_q(\tau)| = o_p(1)$ and $(\phi_1(\cdot),\phi_0(\cdot),\phi_s(\cdot))$ are defined in Section F. Recalling the linear expansion of $\sqrt{n}(\hat{q}^{adj}(\tau) - q(\tau))$ established in Section F, we have

$$
\begin{aligned}
\sqrt{n}(\hat{q}^w(\tau) - \hat{q}^{adj}(\tau)) &= \sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i\in[n]}(\xi_i-1)A_i 1\{S_i=s\}\phi_1(\tau,s,Y_i(1),X_i) \\
&- \sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i\in[n]}(\xi_i-1)(1-A_i)1\{S_i=s\}\phi_0(\tau,s,Y_i(0),X_i) + \frac{1}{\sqrt{n}}\sum_{i\in[n]}(\xi_i-1)\phi_s(\tau,S_i) + R^d_q(\tau) \\
&= W^w_{n,1}(\tau) - W^w_{n,2}(\tau) + R^d_q(\tau),
\end{aligned}
$$

where $\sup_{\tau\in\Upsilon}|R^d_q(\tau)| = o_p(1)$,

$$
W^w_{n,1}(\tau) = \sum_{s\in\mathcal{S}}\frac{1}{\sqrt{n}}\sum_{i\in[n]}(\xi_i-1)A_i 1\{S_i=s\}\phi_1(\tau,s,Y_i(1),X_i)
$$

$$
- \sum_{s \in \mathcal{S}} \frac{1}{\sqrt{n}} \sum_{i \in [n]} (\xi_i - 1)(1 - A_i) 1\{S_i = s\} \phi_0(\tau, s, Y_i(0), X_i),
$$

and

$$
W^w_{n,2}(\tau) = \frac{1}{\sqrt{n}} \sum_{i \in [n]} (\xi_i - 1) \phi_s(\tau, S_i).
$$

Lemma P.5 shows that, uniformly over $\tau \in \Upsilon$ and conditionally on data,

$$
W^w_{n,1}(\tau) + W^w_{n,2}(\tau) \rightsquigarrow \mathcal{B}(\tau),
$$

where $\mathcal{B}(\tau)$ is the Gaussian process with covariance kernel

$$
\begin{aligned}
&\Sigma(\tau, \tau') = \mathbb{E}\pi(S_i)\phi_1(\tau, S_i, Y_i(1), X_i)\phi_1(\tau', S_i, Y_i(1), X_i) \\
&+ \mathbb{E}(1 - \pi(S_i))\phi_0(\tau, S_i, Y_i(0), X_i)\phi_0(\tau', S_i, Y_i(0), X_i) + \mathbb{E}\phi_s(\tau, S_i)\phi_s(\tau', S_i),
\end{aligned}
$$

as defined in Theorem 3.1. This concludes the proof.

H Proof of Theorem 5.1

The proof is divided into two steps. In the first step, we show Assumption 5. Assumption 3(i) can be shown in the same manner and is omitted. In the second step, we establish Assumptions 3(ii) and 3(iii).

Step 1. Recall

$$
\overline{\Delta}_a(\tau, s, X_i) = \hat{m}_a(\tau, s, X_i) - \overline{m}_a(\tau, s, X_i) = \Lambda_{a,s}(X_i, \theta_{a,s}(\tau)) - \Lambda_{a,s}(X_i, \hat{\theta}_{a,s}(\tau)),
$$

and $\{X^s_i, \xi^s_i\}_{i \in [n]}$ is generated independently from the joint distribution of (X_i, ξ_i) given $S_i = s$, and so is independent of $\{A_i, S_i\}_{i \in [n]}$. Let $H_{a,s}(\theta_1, \theta_2) = \mathbb{E}[\Lambda_{a,s}(X_i, \theta_1) - \Lambda_{a,s}(X_i, \theta_2) | S_i = s] = \mathbb{E}[\Lambda_{a,s}(X^s_i, \theta_1) - \Lambda_{a,s}(X^s_i, \theta_2)]$. We have

$$
\begin{aligned}
&\sup_{\tau \in \Upsilon, s \in \mathcal{S}} \left| \frac{\sum_{i \in I_1(s)} \xi_i \overline{\Delta}_1(\tau, s, X_i)}{n^w_1(s)} - \frac{\sum_{i \in I_0(s)} \xi_i \overline{\Delta}_1(\tau, s, X_i)}{n^w_0(s)} \right| \\
&\leq \left(\max_{s \in \mathcal{S}} \frac{n_1(s)}{n^w_1(s)} \right) \sup_{\tau \in \Upsilon, s \in \mathcal{S}} \left| \frac{\sum_{i \in I_1(s)} \xi_i [\overline{\Delta}_1(\tau, s, X_i) - H_{1,s}(\theta_{a,s}(\tau), \hat{\theta}_{a,s}(\tau))]}{n_1(s)} \right| \\
&+ \left(\max_{s \in \mathcal{S}} \frac{n_0(s)}{n^w_0(s)} \right) \sup_{\tau \in \Upsilon, s \in \mathcal{S}} \left| \frac{\sum_{i \in I_0(s)} \xi_i [\overline{\Delta}_1(\tau, s, X_i) - H_{1,s}(\theta_{a,s}(\tau), \hat{\theta}_{a,s}(\tau))]}{n_0(s)} \right| \\
&= o_p(n^{-1/2}).
\end{aligned} \tag{H.1}
$$

To see the last equality, we note that, for any $\varepsilon > 0$, with probability approaching one (w.p.a.1), we have

$$\sup_{\tau\in\Upsilon, s\in\mathcal{S}} ||\hat{\theta}_{a,s}(\tau) - \theta_{a,s}(\tau)|| \leq \varepsilon.$$

Therefore, on the event $\mathcal{A}_n(\varepsilon) \equiv \{\sup_{\tau\in\Upsilon,s\in\mathcal{S}} ||\hat{\theta}_{a,s}(\tau) - \theta_{a,s}(\tau)|| \leq \varepsilon, \min_{s\in\mathcal{S}} n_1(s) \geq \varepsilon n\}$ we have

$$\begin{aligned}
&\sup_{\tau\in\Upsilon}\left|\frac{\sum_{i\in I_1(s)}\xi_i[\overline{\Delta}_1(\tau,s,X_i) - H_{1,s}(\theta_{a,s}(\tau),\hat{\theta}_{a,s}(\tau))]}{n_1(s)}\right|\Bigg|\{A_i,S_i\}_{i\in[n]} \\
&\stackrel{d}{=} \sup_{\tau\in\Upsilon}\left|\frac{\sum_{i=N(s)+1}^{N(s)+n_1(s)}\xi_i^s[\overline{\Delta}_1(\tau,s,X_i^s) - H_{1,s}(\theta_{a,s}(\tau),\hat{\theta}_{a,s}(\tau))]}{n_1(s)}\right|\Bigg|\{A_i,S_i\}_{i\in[n]} \\
&\leq ||\mathbb{P}_{n_1(s)} - \mathbb{P}||_{\mathcal{F}}\Big|\{A_i,S_i\}_{i\in[n]},
\end{aligned}$$

where

$$\mathcal{F} = \{\xi_i^s[\Lambda_{a,s}(X_i^s,\theta_1) - \Lambda_{a,s}(X_i^s,\theta_2) - H_{1,s}(\theta_1,\theta_2)] : ||\theta_1 - \theta_2|| \leq \varepsilon\}.$$

By Assumption 6, $\mathcal{F}$ is a VC-class with a fixed VC index and envelope L_i. In addition,

$$\sup_{f\in\mathcal{F}} \mathbb{P}f^2 \leq \mathbb{E}L_i^2(\theta_1 - \theta_2)^2 \leq C\varepsilon^2.$$

Therefore, for any $\delta > 0$ we have

$$\begin{aligned}
&\mathbb{P}\left(\sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{\sum_{i\in I_1(s)}\xi_i[\overline{\Delta}_1(\tau,s,X_i) - H_{1,s}(\theta_{a,s}(\tau),\hat{\theta}_{a,s}(\tau))]}{n_1(s)}\right| \geq \delta n^{-1/2}\right) \\
&\leq \mathbb{P}\left(\sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{\sum_{i\in I_1(s)}\xi_i[\overline{\Delta}_1(\tau,s,X_i) - H_{1,s}(\theta_{a,s}(\tau),\hat{\theta}_{a,s}(\tau))]}{n_1(s)}\right| \geq \delta n^{-1/2}, \mathcal{A}_n(\varepsilon)\right) + \mathbb{P}(\mathcal{A}_n^c(\varepsilon)) \\
&\leq \mathbb{E}\left[\mathbb{P}\left(\sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{\sum_{i\in I_1(s)}\xi_i[\overline{\Delta}_1(\tau,s,X_i) - H_{1,s}(\theta_{a,s}(\tau),\hat{\theta}_{a,s}(\tau))]}{n_1(s)}\right| \geq \delta n^{-1/2}, \mathcal{A}_n(\varepsilon)\Bigg|\{A_i,S_i\}_{i\in[n]}\right)\right] \\
&+ \mathbb{P}(\mathcal{A}_n^c(\varepsilon)) \\
&\leq \sum_{s\in\mathcal{S}}\mathbb{E}\left[\mathbb{P}\left(||\mathbb{P}_{n_1(s)} - \mathbb{P}||_{\mathcal{F}} \geq \delta n^{-1/2}\Bigg|\{A_i,S_i\}_{i\in[n]}\right)1\{n_1(s) \geq n\varepsilon\}\right] + \mathbb{P}(\mathcal{A}_n^c(\varepsilon)) \\
&\leq \sum_{s\in\mathcal{S}}\mathbb{E}\left\{\frac{n^{1/2}\mathbb{E}\left[||\mathbb{P}_{n_1(s)} - \mathbb{P}||_{\mathcal{F}}|\{A_i,S_i\}_{i\in[n]}\right]1\{n_1(s) \geq n\varepsilon\}}{\delta}\right\} + \mathbb{P}(\mathcal{A}_n^c(\varepsilon)).
\end{aligned}$$

By Chernozhukov, Chetverikov, and Kato (2014, Corollary 5.1),

$$n^{1/2}\mathbb{E}\left[||\mathbb{P}_{n_1(s)} - \mathbb{P}||_{\mathcal{F}}|\{A_i,S_i\}_{i\in[n]}\right]1\{n_1(s) \geq n\varepsilon\}$$

$$
\begin{aligned}
&\leq C(\sqrt{\frac{n}{n_1(s)}\varepsilon^2} + n^{1/2}n_1^{1/q-1}(s))1\{n_1(s) \geq n\varepsilon\} \\
&\leq C(\varepsilon^{1/2} + n^{1/q-1/2}\varepsilon^{1/q-1}).
\end{aligned}
$$

Therefore,

$$
\mathbb{E}\left\{\frac{n^{1/2}\mathbb{E}\left[||\mathbb{P}_{n_1(s)} - \mathbb{P}||_{\mathcal{F}}|\{A_i, S_i\}_{i\in[n]}\right] 1\{n_1(s) \geq n\varepsilon\}}{\delta}\right\} \leq C\mathbb{E}\left(\varepsilon^{1/2} + n^{1/q-1/2}\varepsilon^{1/q-1}\right)/\delta.
$$

By letting $n \to \infty$ followed by $\varepsilon \to 0$, we have

$$
\lim_{n\to\infty}\mathbb{P}\left(\sup_{\tau\in\Upsilon, s\in\mathcal{S}}\left|\frac{\sum_{i\in I_1(s)}\xi_i[\overline{\Delta}_1(\tau, s, X_i) - H_{1,s}(\theta_{a,s}(\tau), \hat{\theta}_{a,s}(\tau))]}{n_1(s)}\right| \geq \delta n^{-1/2}\right) = 0,
$$

In addition,

$$
\max_{s\in\mathcal{S}}|n_1^w(s)/n_1(s) - 1| = \max_{s\in\mathcal{S}}|(D_n^w(s) - D_n(s))/(\pi(s)n(s) + D_n(s))| \xrightarrow{p} 1,
$$

as Lemma P.4 shows that $\max_{s\in\mathcal{S}}|(D_n^w(s) - D_n(s))/n(s)| = o_p(1)$.

Therefore,

$$
\sup_{\tau\in\Upsilon, s\in\mathcal{S}}\left|\frac{\sum_{i\in I_1(s)}\xi_i[\overline{\Delta}_1(\tau, s, X_i) - H_{1,s}(\theta_{a,s}(\tau), \hat{\theta}_{a,s}(\tau))]}{n_1^w(s)}\right| = o_p(n^{-1/2}).
$$

For the same reason, we have

$$
\sup_{\tau\in\Upsilon, s\in\mathcal{S}}\left|\frac{\sum_{i\in I_0(s)}\xi_i[\overline{\Delta}_1(\tau, s, X_i) - H_{1,s}(\theta_{a,s}(\tau), \hat{\theta}_{a,s}(\tau))]}{n_0^w(s)}\right| = o_p(n^{-1/2}),
$$

and (H.1) holds.

Step 2. By Assumption 6,

$$
\begin{aligned}
&|\overline{m}_a(\tau_2, S_i, X_i) - \overline{m}_a(\tau_1, S_i, X_i)| \\
&\leq |\tau_2 - \tau_1| + |\Lambda_{a,s}(X_i, \theta_{a,s}(\tau_1)) - \Lambda_{a,s}(X_i, \theta_{a,s}(\tau_2))| \\
&\leq |\tau_2 - \tau_1| + L_i|\theta_{a,s}(\tau_1) - \theta_{a,s}(\tau_2)| \leq (CL_i + 1)|\tau_2 - \tau_1|.
\end{aligned}
$$

This implies Assumption 3(iii). Furthermore, by Assumption 6 we can let the envelope for the class of functions $\mathcal{F} = \{\overline{m}_a(\tau_2, S_i, X_i) : \tau \in \Upsilon\}$ be $F_i = \max(C, 1)L_i + 1$ where the constant C is the one in the above display. Then, we have

$$
\sup_Q N(\mathcal{F}, e_Q, \varepsilon||F||_{Q,2}) \leq N(\Upsilon, d, \varepsilon) \leq 1/\varepsilon,
$$

where $d(\tau_1, \tau_2) = |\tau_1 - \tau_2|$. This verifies Assumption 3(ii).

I Proof of Theorem 5.2

Recall $\Sigma(\tau,\tau)$ is the asymptotic variance of $\hat{q}^{adj}(\tau)$, and $(\delta_a(\tau, S_i, X_i), \overline{\delta}_a(\tau, S_i, X_i))$ are defined in (F.3). Following the proof of the second part of Theorem 3.1, we have

$$
\begin{aligned}
&\Sigma(\tau,\tau) \\
&= \Bigg\{\mathbb{E}\left[\frac{(\tau - 1\{Y_i(1) \le q_1(\tau)\} - m_1(\tau, S_i, X_i))^2}{\pi(S_i) f_1^2(q_1(0))}\right] + \mathbb{E}\left[\frac{(\tau - 1\{Y_i(0) \le q_0(\tau)\} - m_0(\tau, S_i, X_i))^2}{(1-\pi(S_i)) f_0^2(q_0(0))}\right] \\
&+ \mathbb{E}\left(\frac{m_1(\tau, S_i, X_i)}{f_1(q_1(\tau))} - \frac{m_0(\tau, S_i, X_i)}{f_0(q_0(\tau))}\right)^2\Bigg\} \\
&+ \Bigg\{\mathbb{E}\Bigg[\sqrt{\frac{1-\pi(S_i)}{\pi(S_i)}}\frac{\delta_1(\tau, S_i, X_i) - \overline{\delta}_1(\tau, S_i, X_i)}{f_1(q_1(\tau))} + \sqrt{\frac{\pi(S_i)}{1-\pi(S_i)}}\frac{\delta_0(\tau, S_i, X_i) - \overline{\delta}_0(\tau, S_i, X_i)}{f_0(q_0(\tau))}\Bigg]^2\Bigg\} \\
&= \Bigg\{\mathbb{E}\left[\frac{(\tau - 1\{Y_i(1) \le q_1(\tau)\} - m_1(\tau, S_i, X_i))^2}{\pi(S_i) f_1^2(q_1(0))}\right] + \mathbb{E}\left[\frac{(\tau - 1\{Y_i(0) \le q_0(\tau)\} - m_0(\tau, S_i, X_i))^2}{(1-\pi(S_i)) f_0^2(q_0(0))}\right] \\
&+ \mathbb{E}\left(\frac{m_1(\tau, S_i, X_i)}{f_1(q_1(\tau))} - \frac{m_0(\tau, S_i, X_i)}{f_0(q_0(\tau))}\right)^2\Bigg\} \\
&+ \sum_{s\in\mathcal{S}} p(s)\mathbb{E}\Bigg\{\Bigg[\sqrt{\frac{1-\pi(s)}{\pi(s)}}\frac{m_1(\tau, s, X_i) - m_1(\tau, s) - (\mathbb{E}(\Lambda_{1,s}(X_i, \theta_{1,s}(\tau))|S_i = s) - \Lambda_{1,s}(X_i, \theta_{1,s}(\tau)))}{f_1(q_1(\tau))} \\
&+ \sqrt{\frac{\pi(s)}{1-\pi(s)}}\frac{m_0(\tau, s, X_i) - m_0(\tau, s) - (\mathbb{E}(\Lambda_{0,s}(X_i, \theta_{0,s}(\tau))|S_i = s) - \Lambda_{0,s}(X_i, \theta_{0,s}(\tau)))}{f_0(q_0(\tau))}\Bigg]^2 \Bigg| S_i = s\Bigg\}.
\end{aligned}
$$

In addition, because the three terms in the first curly braces on the RHS of the above display do not depend on $\theta_{a,s}(\tau)$, to minimize $\Sigma(\tau,\tau)$ we only need to minimize the last term. Then, we have

$$
\begin{aligned}
(\theta_{1,s}(\tau), \theta_{0,s}(\tau)) \in \arg\min_{\theta_1,\theta_0} &\mathbb{E}\Bigg\{\Bigg[\sqrt{\frac{1-\pi(s)}{\pi(s)}}\frac{m_1(\tau, s, X_i) - m_1(\tau, s) - g_{1,s}(X_i, \theta_1)}{f_1(q_1(\tau))} \\
&+ \sqrt{\frac{\pi(s)}{1-\pi(s)}}\frac{m_0(\tau, s, X_i) - m_0(\tau, s) - g_{0,s}(X_i, \theta_0)}{f_0(q_0(\tau))}\Bigg]^2 \Bigg| S_i = s\Bigg\} \\
\in \arg\min_{\theta_1,\theta_0} &\mathbb{E}\Bigg\{\left(\frac{g_{1,s}(X_i, \theta_1)}{f_1(q_1(\tau))} + \frac{\pi(s)}{1-\pi(s)}\frac{g_{0,s}(X_i, \theta_0)}{f_0(q_0(\tau))}\right)^2 \\
&- 2\left(\frac{g_{1,s}(X_i, \theta_1)}{f_1(q_1(\tau))} + \frac{\pi(s)}{1-\pi(s)}\frac{g_{0,s}(X_i, \theta_0)}{f_0(q_0(\tau))}\right)\left(\frac{m_1(\tau, s, X_i) - m_1(\tau, s)}{f_1(q_1(\tau))}\right) \\
&- \frac{2\pi(s)}{1-\pi(s)}\left(\frac{g_{1,s}(X_i, \theta_1)}{f_1(q_1(\tau))} + \frac{\pi(s)}{1-\pi(s)}\frac{g_{0,s}(X_i, \theta_0)}{f_0(q_0(\tau))}\right)\left(\frac{m_0(\tau, s, X_i) - m_0(\tau, s)}{f_0(q_0(\tau))}\right) \Bigg| S_i = s\Bigg\},
\end{aligned}
$$

where $g_{a,s}(X_i, \theta_a) = \mathbb{E}(\Lambda_{a,s}(X_i, \theta_a)|S_i = s) - \Lambda_{a,s}(X_i, \theta_a)$.

Next we turn to the second part of Theorem 5.2. Following the previous proof, under the linear probability model with pseudo true values $(\theta_{1,s}(\tau_k), \theta_{0,s}(\tau_k))_{k\in[K]}$, we have

$$\Sigma^{LP}(\tau_k,\tau_l) = \Sigma_1^{LP}(\tau_k,\tau_l) + \sum_{s\in\mathcal{S}} p(s)\mathbb{E}\left[(\tilde{X}_{i,s}\beta_s(\tau_k) - \overline{y}_{i,s}(\tau_k))(\tilde{X}_{i,s}\beta_s(\tau_l) - \overline{y}_{i,s}(\tau_l))|S_i = s\right],$$

where

$$\begin{aligned}
\Sigma_1^{LP}(\tau_k,\tau_l) = &\Bigg\{\mathbb{E}\left[\frac{(\tau - 1\{Y_i(1)\le q_1(\tau)\} - m_1(\tau,S_i,X_i))^2}{\pi(S_i)f_1^2(q_1(0))}\right] \\
&+ \mathbb{E}\left[\frac{(\tau - 1\{Y_i(0)\le q_0(\tau)\} - m_0(\tau,S_i,X_i))^2}{(1-\pi(S_i))f_0^2(q_0(0))}\right] \\
&+ \mathbb{E}\left(\frac{m_1(\tau,S_i,X_i)}{f_1(q_1(\tau))} - \frac{m_0(\tau,S_i,X_i)}{f_0(q_0(\tau))}\right)^2\Bigg\}, \\
\beta_s(\tau) = &\sqrt{\frac{1-\pi(s)}{\pi(s)}}\frac{\theta_{1,s}(\tau)}{f_1(q_1(\tau))} + \sqrt{\frac{\pi(s)}{1-\pi(s)}}\frac{\theta_{0,s}(\tau)}{f_0(q_0(\tau))}, \quad \text{and} \\
\overline{y}_{i,s}(\tau) = &\sqrt{\frac{1-\pi(s)}{\pi(s)}}\frac{[\mathbb{P}(Y_i(1)\le q_1(\tau)|X_i,S_i=s) - \mathbb{P}(Y_i(1)\le q_1(\tau)|S_i=s)]}{f_1(q_1(\tau))} \\
&+ \sqrt{\frac{\pi(s)}{1-\pi(s)}}\frac{[\mathbb{P}(Y_i(0)\le q_0(\tau)|X_i,S_i=s) - \mathbb{P}(Y_i(0)\le q_0(\tau)|S_i=s)]}{f_0(q_0(\tau))}.
\end{aligned}$$

To minimize $[\Sigma^{LP}(\tau_k,\tau_l)]_{k,l\in[K]}$ (in the matrix sense) is the same as minimizing

$$\left[\mathbb{E}\left[(\tilde{X}_{i,s}\beta_s(\tau_k) - \overline{y}_{i,s}(\tau_k))(\tilde{X}_{i,s}\beta_s(\tau_l) - \overline{y}_{i,s}(\tau_l))|S_i = s\right]\right]_{k,l\in[K]}$$

for each $s\in\mathcal{S}$, which is achieved if

$$\beta_s(\tau_k) = [\mathbb{E}\tilde{X}_{i,s}\tilde{X}_{i,s}'|S_i=s]^{-1}\mathbb{E}[\tilde{X}_{i,s}\overline{y}_{i,s}(\tau_k)|S_i=s]. \tag{I.1}$$

Because $\mathbb{E}[\tilde{X}_{i,s}\mathbb{P}(Y_i(a)\le q_a(\tau)|S_i=s)|S_i=s] = 0$ for $a = 0,1$, (I.1) implies

$$\begin{aligned}
&\sqrt{\frac{1-\pi(s)}{\pi(s)}}\frac{\theta_{1,s}(\tau_k)}{f_1(q_1(\tau_k))} + \sqrt{\frac{\pi(s)}{1-\pi(s)}}\frac{\theta_{0,s}(\tau_k)}{f_0(q_0(\tau_k))} \\
&= \sqrt{\frac{1-\pi(s)}{\pi(s)}}\frac{\theta_{1,s}^{LP}(\tau_k)}{f_1(q_1(\tau_k))} + \sqrt{\frac{\pi(s)}{1-\pi(s)}}\frac{\theta_{0,s}^{LP}(\tau_k)}{f_0(q_0(\tau_k))},
\end{aligned}$$

or equivalently,

$$\frac{\theta_{1,s}(\tau_k)}{f_1(q_1(\tau_k))} + \frac{\pi(s)}{1-\pi(s)}\frac{\theta_{0,s}(\tau_k)}{f_0(q_0(\tau_k))} = \frac{\theta_{1,s}^{LP}(\tau_k)}{f_1(q_1(\tau_k))} + \frac{\pi(s)}{1-\pi(s)}\frac{\theta_{0,s}^{LP}(\tau_k)}{f_0(q_0(\tau_k))}.$$

This concludes the proof.

J Proof of Proposition 5.1

By Assumption 8, we see that $\sup_{\tau\in\Upsilon}|\partial_\tau \theta^{LP}_{a,s}(\tau)| < \infty$. This implies Assumption 6(ii). Next, we aim to show

$$\sup_{\tau\in\Upsilon, a=0,1, s\in\mathcal{S}} |\hat{\theta}^{LP}_{a,s}(\tau) - \theta^{LP}_{a,s}(\tau)| = O_p(n^{-1/2}).$$

Focusing on $\hat{\theta}^{LP}_{1,s}(\tau)$ we have

$$\hat{\theta}^{LP}_{1,s}(\tau) - \theta^{LP}_{1,s}(\tau) = \left[\frac{1}{n_1(s)}\sum_{i\in I_1(s)} \hat{X}_{i,1,s}\hat{X}_{i,1,s}^\top\right]^{-1}\left[\frac{1}{n_1(s)}\sum_{i\in I_1(s)} \hat{X}_{i,1,s}(1\{Y_i \le \hat{q}_1(\tau)\} - \hat{X}_{i,1,s}\theta^{LP}_{1,s}(\tau))\right]. \tag{J.1}$$

For the first term in (J.1), we have

$$\frac{1}{n_1(s)}\sum_{i\in I_1(s)} \hat{X}_{i,1,s}\hat{X}_{i,1,s}^\top \stackrel{d}{=} \frac{1}{n_1(s)}\sum_{i=N(s)+1}^{N(s)+n_1(s)} \hat{X}_i^s \hat{X}_i^{s\prime},$$

where $\hat{X}_i^s = X_i^s - \frac{1}{n_1(s)}\sum_{i=N(s)+1}^{N(s)+n_1(s)} X_i^s$ and X_i^s is as defined in Section E. As $\{X_i^s\}_{i=N(s)+1}^{N(s)+n_1(s)}$ is a sequence of i.i.d random variables that is independent of $N(s), n_1(s)$, we have

$$\frac{1}{n_1(s)}\sum_{i=N(s)+1}^{N(s)+n_1(s)} \hat{X}_i^s\hat{X}_i^{s\top} \stackrel{p}{\longrightarrow} \mathbb{E}(X_i^s - \mathbb{E}X_i^s)(X_i^s - \mathbb{E}X_i^s)^\top = \mathbb{E}(\tilde{X}_{i,s}\tilde{X}_{i,s}^\top|S_i = s),$$

where $\tilde{X}_{i,s} = X_i - \mathbb{E}(X_i|S_i = s)$. For the second term in (J.1), we have

$$\begin{aligned}
&\frac{1}{n_1(s)}\sum_{i\in I_1(s)} \hat{X}_{i,1,s}\left(1\{Y_i \le \hat{q}_1(\tau)\} - \hat{X}_{i,1,s}\theta^{LP}_{1,s}(\tau)\right) \\
&= \frac{1}{n_1(s)}\sum_{i\in I_1(s)} \tilde{X}_{i,s}\left(1\{Y_i \le \hat{q}_1(\tau)\} - \hat{X}_{i,1,s}\theta^{LP}_{1,s}(\tau)\right) + R_1(\tau) \\
&= \frac{1}{n_1(s)}\sum_{i\in I_1(s)} \tilde{X}_{i,s}\left(1\{Y_i \le \hat{q}_1(\tau)\} - \tilde{X}_{i,s}\theta^{LP}_{1,s}(\tau)\right) + R_1(\tau) + R_2(\tau) \\
&= \left[\frac{1}{n_1(s)}\sum_{i\in I_1(s)} \tilde{X}_{i,s}(1\{Y_i \le \hat{q}_1(\tau)\} - 1\{Y_i \le q_1(\tau)\})\right] \\
&+ \left[\frac{1}{n_1(s)}\sum_{i\in I_1(s)} \tilde{X}_{i,s}(1\{Y_i \le q_1(\tau)\} - \tilde{X}_{i,s}\theta^{LP}_{1,s}(\tau))\right] + R_1(\tau) + R_2(\tau) \\
&\equiv I(\tau) + II(\tau) + R_1(\tau) + R_2(\tau),
\end{aligned}$$

where

$$R_1(\tau) = -\left(\frac{1}{n_1(s)}\sum_{i\in I_1(s)} X_i - \mathbb{E}(X_i|S_i=s)\right)\left(\frac{1}{n_1(s)}\sum_{i\in I_1(s)}(1\{Y_i \le \hat{q}_1(\tau)\} - \hat{X}_{i,1,s}\theta^{LP}_{1,s}(\tau))\right),$$

and

$$R_2(\tau) = \left(\frac{1}{n_1(s)}\sum_{i\in I_1(s)} X_i - \mathbb{E}(X_i|S_i=s)\right)\left(\frac{1}{n_1(s)}\sum_{i\in I_1(s)}\tilde{X}_{i,s}\theta^{LP}_{1,s}(\tau)\right).$$

By Assumption 8 we can show that $\sup_{\tau\in\Upsilon}|\theta^{LP}_{1,s}(\tau)| \le C < \infty$ for some constant $C > 0$. Therefore, we have

$$\begin{aligned}\sup_{\tau\in\Upsilon}|R_1(\tau)| &= \sup_{\tau\in\Upsilon}\left|\frac{1}{n_1(s)}\sum_{i\in I_1(s)} X_i - \mathbb{E}(X_i|S_i=s)\right|\left|\frac{1}{n_1(s)}\sum_{i\in I_1(s)}(1\{Y_i \le \hat{q}_1(\tau)\} - \hat{X}_{i,1,s}\theta^{LP}_{1,s}(\tau))\right| \\ &= O_p(n^{-1/2}),\end{aligned}$$

and

$$\begin{aligned}\sup_{\tau\in\Upsilon}|R_2(\tau)| &= \sup_{\tau\in\Upsilon}\left|\frac{1}{n_1(s)}\sum_{i\in I_1(s)} X_i - \mathbb{E}(X_i|S_i=s)\right|\left|\frac{1}{n_1(s)}\sum_{i\in I_1(s)}\tilde{X}_{i,s}\theta^{LP}_{1,s}(\tau)\right| \\ &= O_p(n^{-1/2}),\end{aligned}$$

where we use the fact that

$$\left|\frac{1}{n_1(s)}\sum_{i\in I_1(s)} X_i - \mathbb{E}(X_i|S_i=s)\right| \stackrel{d}{=} \left|\frac{1}{n_1(s)}\sum_{i=N(s)+1}^{N(s)+n_1(s)}(X_i^s - \mathbb{E}X_i^s)\right| = O_p(n^{-1/2}).$$

Next, note that $\sup_{\tau\in\Upsilon}|\hat{q}(\tau) - q(\tau)| = O_p(n^{-1/2})$, which means for any $\varepsilon > 0$, there exists a constant $M > 0$ such that $\sup_{\tau\in\Upsilon}|\hat{q}(\tau) - q(\tau)| \le Mn^{-1/2}$ with probability greater than $1-\varepsilon$. On the event set that $\sup_{\tau\in\Upsilon}|\hat{q}(\tau) - q(\tau)| \le Mn^{-1/2}$, we have

$$\begin{aligned}\sup_{\tau\in\Upsilon}|I(\tau)| \le \sup_{\tau\in\Upsilon}&\left|\frac{1}{n_1(s)}\sum_{i\in I_1(s)}\tilde{X}_{i,s}\Big(1\{Y_i(1) \le \hat{q}_1(\tau)\} - 1\{Y_i(1) \le q_1(\tau)\}\right. \\ &\left. - \mathbb{P}(Y_i(1) \le \hat{q}_1(\tau)|X_i, S_i=s) + \mathbb{P}(Y_i(1) \le q_1(\tau)|X_i, S_i=s)\Big)\right| \\ &+ \sup_{\tau\in\Upsilon}\left|\frac{1}{n_1(s)}\sum_{i\in I_1(s)}\tilde{X}_{i,s}(\mathbb{P}(Y_i(1) \le \hat{q}_1(\tau)|X_i, S_i=s) - \mathbb{P}(Y_i(1) \le q_1(\tau)|X_i, S_i=s))\right| \\ &\le \sup_{|q-q'|\le Mn^{-1/2}}\left|\frac{1}{n_1(s)}\sum_{i\in I_1(s)}\tilde{X}_{i,s}\Big(1\{Y_i(1) \le q\} - 1\{Y_i(1) \le q'\}\right.\end{aligned}$$

$$
\begin{aligned}
& - \mathbb{P}(Y_i(1) \leq q | X_i, S_i = s) + \mathbb{P}(Y_i(1) \leq q' | X_i, S_i = s)\Big)\Bigg| + C \sup_{\tau \in \Upsilon} |\hat{q}(\tau) - q(\tau)| \\
& \leq \sup_{|q-q'| \leq Mn^{-1/2}} \Bigg| \frac{1}{n_1(s)} \sum_{i \in I_1(s)} \tilde{X}_{i,s} \Big(1\{Y_i(1) \leq q\} - 1\{Y_i(1) \leq q'\} \\
& - \mathbb{P}(Y_i(1) \leq q | X_i, S_i = s) + \mathbb{P}(Y_i(1) \leq q' | X_i, S_i = s)\Big)\Bigg| + Cn^{-1/2} \\
& = O_p(n^{-1/2}),
\end{aligned}
$$

where the first inequality is due to the triangle inequality, the second inequality is due to the fact that $\sup_{\tau \in \Upsilon} |\hat{q}(\tau) - q(\tau)| \leq Mn^{-1/2}$, and the third inequality is due to the fact that $f_1(\cdot|X_i, S_i = s)$ is assumed to be bounded. To see the last equality in the above display, we define

$$
\mathcal{F} = \left\{ \begin{array}{c} (X_{i,j} - \mathbb{E}X_{i,j}|S_i = s)\Big(1\{Y_i(1) \leq q\} - 1\{Y_i(1) \leq q'\} \\ -\mathbb{P}(Y_i(1) \leq q | X_i, S_i = s) + \mathbb{P}(Y_i(1) \leq q' | X_i, S_i = s)\Big) : |q - q'| \leq Mn^{-1/2} \end{array} \right\}
$$

with envelope $F_i = 2|X_{i,j} - \mathbb{E}(X_{i,j}|S_i = s)| \in L_{\mathbb{P},q}$ for some $q > 2$, where $X_{i,j}$ is the j-th coordinate of X_i. Clearly $\mathcal{F}$ is of the VC-type with fixed coefficients (α, v). In addition,

$$
\sup_{f \in \mathcal{F}} \mathbb{P}f^2 \leq Cn^{-1/2} \equiv \sigma_n^2.
$$

Therefore, Lemma P.2 implies that $\sup_{\tau \in \Upsilon} |I(\tau)| = O_p(n^{-1/2})$. By the usual maximal inequality (e.g. van der Vaart and Wellner, 1996, Theorem 2.14.1), we can show thjat

$$
\sup_{\tau \in \Upsilon} |II(\tau)| = O_p(n^{-1/2}).
$$

Combining these results, we conclude that

$$
\sup_{\tau \in \Upsilon, s \in \mathcal{S}} \left| \frac{1}{n_1(s)} \sum_{i \in I_1(s)} \hat{X}_{i,1,s}(1\{Y_i \leq \hat{q}_1(\tau)\} - \hat{X}_{i,1,s}\theta_{1,s}^{LP}(\tau)) \right| = O_p(n^{-1/2}),
$$

and hence

$$
\sup_{\tau \in \Upsilon, s \in \mathcal{S}} |\hat{\theta}_{1,s}^{LP}(\tau) - \theta_{1,s}^{LP}(\tau)| = O_p(n^{-1/2}).
$$

K Proof of Proposition 5.2

Note $\hat{\theta}^{LG}_{a,s}(\tau) \in \Theta_a$, which is a compact subset of $\Re^{d_x}$. In addition, by Assumption 9

$$\sup_{\tau \in \Upsilon} ||\theta^{LG}_{a,s}(\tau)||_2 < \infty.$$

Therefore, there exists a constant $M > 0$ such that

$$\sup_{\tau \in \Upsilon} ||\hat{\theta}^{LG}_{a,s}(\tau) - \theta^{LG}_{a,s}(\tau)||_2 \leq M.$$

Recall $Q(\tau, s, \theta_1, \theta_0)$ defined in Theorem 5.2 with $\Lambda_{a,s}(X_i, \theta_a) = \lambda(\vec{X}_i^\top \theta_a)$. For some $\delta_0 > 0$, let $\mathcal{D} = \{(\delta_1, \delta_0) \in \Re^{2dx}, \sqrt{||\delta_1||_2^2 + ||\delta_0||_2^2} \in [\delta_0, M]\}$, which is compact and

$$\eta = \inf_{(\tau,\delta) \in \Upsilon \times \mathcal{D}} (Q(s, \tau, \theta^{LG}_{1,s}(\tau) + \delta_1, \theta^{LG}_{0,s}(\tau) + \delta_0) - Q(s, \tau, \theta^{LG}_{1,s}(\tau), \theta^{LG}_{0,s}(\tau))).$$

Because $Q(s, \tau, \theta^{LG}_{1,s}(\tau) + \delta_1, \theta^{LG}_{0,s}(\tau) + \delta_0) - Q(s, \tau, \theta^{LG}_{1,s}(\tau), \theta^{LG}_{0,s}(\tau))$ is a continuous function of (δ, τ), $\Upsilon \times \mathcal{D}$ is compact, and $(\theta^{LG}_{1,s}(\tau), \theta^{LG}_{0,s}(\tau))$ is the unique minimizer of $Q(s, \tau, \theta_1, \theta_0)$, we have $\eta > 0$. On the other hand

$$\begin{aligned}
&(Q(s, \tau, \theta^{LG}_{1,s}(\tau) + \delta_1, \theta^{LG}_{0,s}(\tau) + \delta_0) - Q(s, \tau, \theta^{LG}_{1,s}(\tau), \theta^{LG}_{0,s}(\tau))) \\
&\leq (Q_n(s, \tau, \theta^{LG}_{1,s}(\tau) + \delta_1, \theta^{LG}_{0,s}(\tau) + \delta_0) - Q_n(s, \tau, \theta^{LG}_{1,s}(\tau), \theta^{LG}_{0,s}(\tau))) + 2\Delta_n,
\end{aligned}$$

where

$$\Delta_n = \sup_{\tau \in \Upsilon, s \in \mathcal{S}, \theta_1, \theta_0 \in \Theta} |Q_n(s, \tau, \theta_1, \theta_0) - Q(s, \tau, \theta_1, \theta_0)|.$$

If $\sup_{\tau \in \Upsilon} \sqrt{||\hat{\theta}^{LG}_{1,s}(\tau) - \theta^{LG}_{1,s}(\tau)||_2^2 + ||\hat{\theta}^{LG}_{0,s}(\tau) - \theta^{LG}_{0,s}(\tau)||_2^2} \geq \delta_0$, then there exists some $\tau \in \Upsilon$ and $\delta_1 = \hat{\theta}^{LG}_{1,s}(\tau) - \theta^{LG}_{1,s}(\tau)$, and $\delta_0 = \hat{\theta}^{LG}_{0,s}(\tau) - \theta^{LG}_{0,s}(\tau)$ such that

$$Q_n(s, \tau, \theta^{LG}_{1,s}(\tau) + \delta_1, \theta^{LG}_{0,s}(\tau) + \delta_0) - Q_n(s, \tau, \theta^{LG}_{1,s}(\tau), \theta^{LG}_{1,s}(\tau)) \leq 0.$$

This implies

$$\mathbb{P}(\sup_{\tau \in \Upsilon} \sqrt{||\hat{\theta}^{LG}_{1,s}(\tau) - \theta^{LG}_{1,s}(\tau)||_2^2 + ||\hat{\theta}^{LG}_{0,s}(\tau) - \theta^{LG}_{0,s}(\tau)||_2^2} \geq \delta_0) \leq \mathbb{P}(\eta \leq 2\Delta_n) \to 0,$$

where the last step holds because Lemma P.6 has established that $\Delta_n = o_p(1)$. As δ_0 is arbitrary, we have

$$\sup_{\tau \in \Upsilon} ||\hat{\theta}^{LG}_{1,s}(\tau) - \theta^{LG}_{1,s}(\tau)||_2 = o_p(1) \quad \text{and} \quad \sup_{\tau \in \Upsilon} ||\hat{\theta}^{LG}_{0,s}(\tau) - \theta^{LG}_{0,s}(\tau)||_2 = o_p(1).$$

L Proof of Proposition 5.3

Let

$$Q_n(\tau, s, q, \theta_a) = \frac{1}{n_a(s)} \sum_{i \in I_a(s)} [1\{Y_i \leq q\} \log(\lambda(H_i^\top \theta_a)) + 1\{Y_i > q\} \log(1 - \lambda(H_i^\top \theta_a))],$$

and

$$Q(\tau, s, q, \theta_a) = \mathbb{E}[1\{Y_i(a) \leq q\} \log(\lambda(H_i^\top \theta_a)) + 1\{Y_i(a) > q\} \log(1 - \lambda(H_i^\top \theta_a))|S_i = s].$$

Following the same argument in the proof of Lemma P.6 (replacing $\vec{X}_i$ by H_i), we can show

$$\sup_{\tau \in \Upsilon, s \in \mathcal{S}, q \in \Re, \theta_a \in \Re^{d_x}} |Q_n(\tau, s, q, \theta_a) - Q(\tau, s, q, \theta_a)| = o_p(1).$$

In addition,

$$\sup_{\tau \in \Upsilon, s \in \mathcal{S}, q \in \Re, \theta_a \in \Re^{d_x}} |\partial_q Q(\tau, s, q, \theta_a)| \leq C,$$

and $\sup_{\tau \in \Upsilon} |\hat{q}_a(\tau) - q_a(\tau)| = O_p(n^{-1/2})$. Therefore,

$$\begin{aligned}
\Delta_n \equiv & \sup_{\tau \in \Upsilon, s \in \mathcal{S}, \theta_a \in \Re^{d_x}} |Q_n(\tau, s, \hat{q}_a(\tau), \theta_a) - Q(\tau, s, q_a(\tau), \theta_a)| \\
\leq & \sup_{\tau \in \Upsilon, s \in \mathcal{S}, q \in \Re, \theta_a \in \Re^{d_x}} |Q_n(\tau, s, q, \theta_a) - Q(\tau, s, q, \theta_a)| \\
+ & \sup_{\tau \in \Upsilon, s \in \mathcal{S}, \theta_a \in \Re^{d_x}} |Q(\tau, s, \hat{q}_a(\tau), \theta_a) - Q(\tau, s, q_a(\tau), \theta_a)| = o_p(1).
\end{aligned} \tag{L.1}$$

In addition, note that $Q_n(\tau, s, \hat{q}_a(\tau), \theta_a)$ is concave in θ_a for fixed τ. Therefore, for $u \in S^{d_x - 1}$ and $l > \delta$

$$Q_n(\tau, s, \hat{q}_a(\tau), \theta_{a,s}^{ML}(\tau) + \delta u) \geq \frac{\delta}{l} Q_n(\tau, s, \hat{q}_a(\tau), \theta_{a,s}^{ML}(\tau) + lu) + (1 - \frac{\delta}{l}) Q_n(\tau, s, \hat{q}_a(\tau), \theta_{a,s}^{ML}(\tau)),$$

which implies

$$\begin{aligned}
& \frac{\delta}{l} \left(Q_n(\tau, s, \hat{q}_a(\tau), \theta_{a,s}^{ML}(\tau) + lu) - Q_n(\tau, s, \hat{q}_a(\tau), \theta_{a,s}^{ML}(\tau)) \right) \\
& \leq Q_n(\tau, s, \hat{q}_a(\tau), \theta_{a,s}^{ML}(\tau) + \delta u) - Q_n(\tau, s, \hat{q}_a(\tau), \theta_{a,s}^{ML}(\tau)) \\
& \leq Q(\tau, s, q_a(\tau), \theta_{a,s}^{ML}(\tau) + \delta u) - Q(\tau, s, q_a(\tau), \theta_{a,s}^{ML}(\tau)) + 2\Delta_n.
\end{aligned}$$

Because $Q(\tau, s, q_a(\tau), \theta_{a,s}^{ML}(\tau) + \delta u) - Q(\tau, s, q_a(\tau), \theta_{a,s}^{ML}(\tau))$ is continuous in $(\tau, u) \in \Upsilon \times S^{d_x - 1}$,

$\Upsilon \times S^{d_x-1}$ is compact, and $\theta^{ML}_{a,s}(\tau)$ is the unique maximizer of $Q(\tau, s, q_a(\tau), \theta_a)$, we have

$$\sup_{(\tau,u)\in\Upsilon\times S^{d_x-1}} Q(\tau, s, q_a(\tau), \theta^{ML}_{a,s}(\tau) + \delta u) - Q(\tau, s, q_a(\tau), \theta^{ML}_{a,s}(\tau)) \leq -\eta,$$

for some $\eta > 0$. In addition, if $\sup_{\tau\in\Upsilon} ||\hat{\theta}^{ML}_{a,s}(\tau) - \theta^{ML}_{a,s}(\tau)||_2 > \delta$, then there exists $(\tau, l, u) \in \Upsilon \times (\delta, \infty) \times S^{d_x-1}$ such that

$$\frac{\delta}{l}(Q_n(\tau, s, \hat{q}_a(\tau), \theta^{ML}_{a,s}(\tau) + lu) - Q_n(\tau, s, \hat{q}_a(\tau), \theta^{ML}_{a,s}(\tau))) \geq 0.$$

Therefore,

$$\mathbb{P}\left(\sup_{\tau\in\Upsilon} ||\hat{\theta}^{ML}_{a,s}(\tau) - \theta^{ML}_{a,s}(\tau)||_2 > \delta\right) \leq \mathbb{P}(\eta \leq 2\Delta_n) \to 0,$$

where the last step is due to (L.1). This implies

$$\sup_{\tau\in\Upsilon} ||\hat{\theta}^{ML}_{a,s}(\tau) - \theta^{ML}_{a,s}(\tau)||_2 = o_p(1).$$

M Proof of Theorem 5.4

The proof strategy follows Belloni, Chernozhukov, Fernández-Val, and Hansen (2017). We provide details here just for completeness. We divide the proof into three steps. In the first step, we show

$$\sup_{\tau\in\Upsilon} ||\hat{\theta}^{NP}_{a,s}(\tau) - \theta^{NP}_{a,s}(\tau)||_2 = O_p(\sqrt{h_n \log(n)/n}).$$

In the second step, we establish Assumption 5. By a similar argument, we can establish Assumption 3(i). In the third step, we establish Assumptions 3(ii) and 3(iii).

Step 1. Let $\hat{U}_\tau = \hat{\theta}^{NP}_{a,s}(\tau) - \theta^{NP}_{a,s}(\tau)$,

$$\begin{aligned} Q_n(\tau, s, q, \theta) &= \frac{-1}{n_a(s)} \sum_{i\in I_a(s)} [1\{Y_i \leq q\} \log(\lambda(H^\top_{h_n}(X_i)\theta_a)) + 1\{Y_i > q\} \log(1 - \lambda(H^\top_{h_n}(X_i)\theta_a))] \\ &= \frac{1}{n_a(s)} \sum_{i\in I_a(s)} [\log\left(1 + \exp(H^\top_{h_n}(X_i)\theta_a)\right) - 1\{Y_i \leq q\} H^\top_{h_n}(X_i)\theta_a], \end{aligned}$$

and for an arbitrary $U_\tau \in \Re^{h_n}$,

$$\ell_i(t) = \log(1 + \exp(H^\top_{h_n}(X_i)(\theta^{NP}_{a,s}(\tau) + tU_\tau))).$$

Then, we have

$$\hat{U}_\tau = \arg\max_{U_\tau} Q_n(\tau, s, \hat{q}_a(\tau), \theta_{a,s}^{NP}(\tau) + U_\tau) - Q_n(\tau, s, \hat{q}_a(\tau), \theta_{a,s}^{NP}(\tau)),$$

$$\partial_t Q_n(\tau, s, \hat{q}_a(\tau), \theta_{a,s}^{NP}(\tau) + tU_\tau)|_{t=0} = \frac{-1}{n_a(s)} \sum_{i \in I_a(s)} \left(1\{Y_i \leq \hat{q}_a(\tau)\} - \lambda(H_{h_n}^\top(X_i)\theta_{a,s}^{NP}(\tau))\right) H_{h_n}(X_i),$$

and

$$\begin{aligned}
&Q_n(\tau, s, \hat{q}_a(\tau), \theta_{a,s}^{NP}(\tau) + U_\tau) - Q_n(\tau, s, \hat{q}_a(\tau), \theta_{a,s}^{NP}(\tau)) - \partial_t Q_n^\top(\tau, s, \hat{q}_a(\tau), \theta_{a,s}^{NP}(\tau) + tU_\tau)|_{t=0} U_\tau \\
=&\frac{1}{n_a(s)} \sum_{i \in I_a(s)} \left[\ell_i(1) - \ell_i(0) - \ell_i'(0)\right].
\end{aligned}$$

In addition

$$|\ell_i'''(t)| \leq |\ell_i''(t)||H_{h_n}^\top(X_i)U_\tau|.$$

Therefore, there exists a constant $\underline{c} > 0$ such that

$$\begin{aligned}
&\ell_i(1) - \ell_i(0) - \ell_i'(0) \\
&\geq \frac{\ell_i''(0)}{(H_{h_n}^\top(X_i)U_\tau)^2} \left[\exp(-|H_{h_n}^\top(X_i)U_\tau|) + |H_{h_n}^\top(X_i)U_\tau| - 1\right] \\
&= \lambda(H_{h_n}^\top(X_i)\theta_{a,s}^{NP}(\tau))(1 - \Lambda(H_{h_n}^\top(X_i)\theta_{a,s}^{NP}(\tau))) \left[\exp(-|H_{h_n}^\top(X_i)U_\tau|) + |H_{h_n}^\top(X_i)U_\tau| - 1\right] \\
&\geq \underline{c} \left[\exp(-|H_{h_n}^\top(X_i)U_\tau|) + |H_{h_n}^\top(X_i)U_\tau| - 1\right] \\
&\geq \underline{c} \left(\frac{(H_{h_n}^\top(X_i)U_\tau)^2}{2} - \frac{|H_{h_n}^\top(X_i)U_\tau|^3}{6}\right),
\end{aligned}$$

where the first inequality is due to Bach (2010, Lemma 1) and the third inequality holds because

$$e^{-x} + x - 1 \geq \frac{x^2}{2} - \frac{x^3}{6},\ x > 0.$$

To see the second inequality, note that $e^{-x} + x - 1 \geq 0$ for $x \geq 0$ and by Assumption 11,

$$\begin{aligned}
&\inf_{a=0,1, s \in \mathcal{S}, \tau \in \Upsilon, x \in \text{Supp}(X)} \lambda(H_{h_n}^\top(x)\theta_{a,s}^{NP}(\tau)) \\
&= \inf_{a=0,1, s \in \mathcal{S}, \tau \in \Upsilon, x \in \text{Supp}(X)} (\mathbb{P}(Y_i(a) \leq q_a(\tau)|S_i = s, X_i = x) - R_a(\tau, s, x)) \geq c/2,
\end{aligned}$$

and

$$
\begin{aligned}
&\sup_{a=0,1,s\in\mathcal{S},\tau\in\Upsilon,x\in\text{Supp}(X)} \lambda(H_{h_n}^\top(x)\theta_{a,s}^{NP}(\tau)) \\
=&\sup_{a=0,1,s\in\mathcal{S},\tau\in\Upsilon,x\in\text{Supp}(X)} (\mathbb{P}(Y_i(a)\leq q_a(\tau)|S_i=s,X_i=x)+R_a(\tau,s,x))\leq 1-c/2.
\end{aligned}
$$

This implies

$$
\inf_{\tau\in\Upsilon}\lambda(H_{h_n}^\top(X_i)\theta_{a,s}^{NP}(\tau))(1-\Lambda(H_{h_n}^\top(X_i)\theta_{a,s}^{NP}(\tau)))\geq \underline{c}>0,
$$

and thus,

$$
\begin{aligned}
G_n(U_\tau)\equiv\, & Q_n(\tau,s,\hat{q}_a(\tau),\theta_{a,s}^{NP}(\tau)+U_\tau)-Q_n(\tau,s,\hat{q}_a(\tau),\theta_{a,s}^{NP}(\tau))-\partial_t Q_n^\top(\tau,s,\hat{q}_a(\tau),\theta_{a,s}^{NP}(\tau)+tU_\tau)|_{t=0}U_\tau \\
\geq\, & \frac{\underline{c}}{n_a(s)}\sum_{i\in I_a(s)}\left(\frac{(H_{h_n}^\top(X_i)U_\tau)^2}{2}-\frac{|H_{h_n}^\top(X_i)U_\tau|^3}{6}\right).
\end{aligned}
$$

Let

$$
\overline{\ell}=\inf_{U\in\Re^{h_n}}\frac{\left[\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(H_{h_n}^\top(X_i)U)^2\right]^{3/2}}{\frac{1}{n_a(s)}\sum_{i\in I_a(s)}|H_{h_n}^\top(X_i)U|^3}. \tag{M.1}
$$

If $\left[\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(H_{h_n}^\top(X_i)\hat{U}_\tau)^2\right]^{1/2}\leq\overline{\ell}$, then

$$
\begin{aligned}
&\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(H_{h_n}^\top(X_i)\hat{U}_\tau)^2 \\
=&\left[\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(H_{h_n}^\top(X_i)\hat{U}_\tau)^2\right]^{-1/2}\frac{\left[\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(H_{h_n}^\top(X_i)\hat{U}_\tau)^2\right]^{3/2}}{\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(H_{h_n}^\top(X_i)\hat{U}_\tau)^3}\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(H_{h_n}^\top(X_i)\hat{U}_\tau)^3 \\
\geq&\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(H_{h_n}^\top(X_i)\hat{U}_\tau)^3,
\end{aligned}
$$

and thus

$$
G_n(\hat{U}_\tau)\geq\frac{\underline{c}}{n_a(s)}\sum_{i\in I_a(s)}\left(\frac{(H_{h_n}^\top(X_i)U_\tau)^2}{2}-\frac{|H_{h_n}^\top(X_i)U_\tau|^3}{6}\right)\geq\frac{\underline{c}}{n_a(s)}\sum_{i\in I_a(s)}\frac{(H_{h_n}^\top(X_i)\hat{U}_\tau)^2}{3}.
$$

On the other hand, if $\left[\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(H_{h_n}^\top(X_i)\hat{U}_\tau)^2\right]^{1/2}>\overline{\ell}$, we can denote $\overline{U}_\tau=\frac{\overline{\ell}\hat{U}_\tau}{\left[\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(H_{h_n}^\top(X_i)\hat{U}_\tau)^2\right]^{1/2}}$

such that

$$\left[\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(H_{h_n}^\top(X_i)\overline{U}_\tau)^2\right]^{1/2}\leq \overline{\ell}.$$

Further, because $G_n(U_\tau)$ is convex in U_τ we have

$$\begin{aligned}
G_n(\hat{U}_\tau) &= G_n\left(\frac{\left[\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(H_{h_n}^\top(X_i)\hat{U}_\tau)^2\right]^{1/2}}{\overline{\ell}}\overline{U}_\tau\right) \\
&\geq \frac{\left[\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(H_{h_n}^\top(X_i)\hat{U}_\tau)^2\right]^{1/2}}{\overline{\ell}}G_n(\overline{U}_\tau) \\
&\geq \frac{\left[\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(H_{h_n}^\top(X_i)\hat{U}_\tau)^2\right]^{1/2}}{\overline{\ell}}\frac{\underline{c}}{n_a(s)}\sum_{i\in I_a(s)}\frac{(H_{h_n}^\top(X_i)\overline{U}_\tau)^2}{3} \\
&= \frac{\underline{c}\overline{\ell}}{3}\left[\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(H_{h_n}^\top(X_i)\hat{U}_\tau)^2\right]^{1/2}.
\end{aligned}$$

Therefore, for some constant $\overline{c}$ that only depends on $\underline{c}$ and κ_1, we have

$$\begin{aligned}
G_n(\hat{U}_\tau) &\geq \min\left(\frac{\underline{c}}{n_a(s)}\sum_{i\in I_a(s)}\frac{(H_{h_n}^\top(X_i)\hat{U}_\tau)^2}{3}, \frac{\underline{c}\overline{\ell}}{3}\left[\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(H_{h_n}^\top(X_i)\hat{U}_\tau)^2\right]^{1/2}\right) \\
&\geq \frac{\overline{c}}{3}\min(||\hat{U}_\tau||_2^2, \overline{\ell}||\hat{U}_\tau||_2). \qquad \text{(M.2)}
\end{aligned}$$

In addition, by construction,

$$\begin{aligned}
G_n(\hat{U}_\tau) &\leq |\partial_t Q_n^\top(\tau, s, \hat{q}_a(\tau), \theta_{a,s}^{NP}(\tau)+tU_\tau)|_{t=0}\hat{U}_\tau| \\
&= \left|\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(1\{Y_i\leq\hat{q}_a(\tau)\}-\lambda(H_{h_n}^\top(X_i)\theta_{a,s}^{NP}(\tau)))H_{h_n}^\top(X_i)\hat{U}_\tau\right| \\
&\leq \left\|\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(1\{Y_i\leq\hat{q}_a(\tau)\}-m_a(\tau,s,X_i))H_{h_n}(X_i)\right\|_\infty \left\|\hat{U}_\tau\right\|_1 \\
&+ \left[\frac{1}{n_a(s)}\sum_{i\in I_a(s)}R_a^2(\tau,s,X_i)\right]^{1/2}\left[\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(H_{h_n}^\top(X_i)\hat{U}_\tau)\right]^{1/2} \\
&\leq \left\|\frac{h_n^{1/2}}{n_a(s)}\sum_{i\in I_a(s)}(1\{Y_i\leq\hat{q}_a(\tau)\}-m_a(\tau,s,X_i))H_{h_n}(X_i)\right\|_\infty \left\|\hat{U}_\tau\right\|_2
\end{aligned}$$

$$+\left[\frac{\kappa_2}{n_a(s)}\sum_{i\in I_a(s)} R_a^2(\tau,s,X_i)\right]^{1/2} ||\hat{U}_\tau||_2. \tag{M.3}$$

Combining (M.2) and (M.3), we have

$$\frac{\overline{c}}{3}\min(||\hat{U}_\tau||_2,\overline{\ell}) \leq \left\|\frac{h_n^{1/2}}{n_a(s)}\sum_{i\in I_a(s)}(1\{Y_i\leq \hat{q}_a(\tau)\}-m_a(\tau,s,X_i))H_{h_n}(X_i)\right\|_\infty + \left[\frac{\kappa_2}{n_a(s)}\sum_{i\in I_a(s)} R_a^2(\tau,s,X_i)\right]^{1/2}.$$

Taking $\sup_{\tau\in\Upsilon}$ on both sides, we have

$$\begin{aligned}
&\frac{\overline{c}}{3}\min(\sup_{\tau\in\Upsilon}||\hat{U}_\tau||_2,\overline{\ell})\\
&\leq \sup_{\tau\in\Upsilon}\left\|\frac{h_n^{1/2}}{n_a(s)}\sum_{i\in I_a(s)}(1\{Y_i\leq \hat{q}_a(\tau)\}-m_a(\tau,s,X_i))H_{h_n}(X_i)\right\|_\infty + \sup_{\tau\in\Upsilon}\left[\frac{\kappa_2}{n_a(s)}\sum_{i\in I_a(s)} R_a^2(\tau,s,X_i)\right]^{1/2}\\
&= O_p(\sqrt{\frac{h_n\log(n)}{n}}),
\end{aligned}$$

where the last line holds due to Assumption 11 and Lemma P.7. Finally, Lemma P.8 shows that $\overline{\ell}/\sqrt{\frac{h_n\log(n)}{n}}\to\infty$, which implies

$$\sup_{\tau\in\Upsilon}||\hat{U}_\tau||_2 = O_p\left(\sqrt{\frac{h_n\log(n)}{n}}\right).$$

Step 2. Recall

$$\begin{aligned}
\overline{\Delta}_a(\tau,s,X_i) &= \hat{m}_a(\tau,s,X_i)-\overline{m}_a(\tau,s,X_i)\\
&= \mathbb{P}(Y_i(a)\leq q_a(\tau)|X_i,S_i=s)-\lambda(H_{h_n}^\top(X_i)\hat{\theta}_{a,s}^{NP}(\tau))\\
&= \lambda(H_{h_n}^\top(X_i)\theta_{a,s}^{NP}(\tau))-\lambda(H_{h_n}^\top(X_i)\hat{\theta}_{a,s}^{NP}(\tau))+R_a(\tau,s,X_i),
\end{aligned}$$

and $\{X_i^s,\xi_i^s\}_{i\in[n]}$ is generated independently from the joint distribution of (X_i,ξ_i) given $S_i=s$, and so is independent of $\{A_i,S_i\}_{i\in[n]}$. Let

$$H(\theta_1,\theta_2)=\mathbb{E}[\lambda(H_{h_n}^\top(X_i)\theta_1)-\lambda(H_{h_n}^\top(X_i)\theta_2)|S_i=s]=\mathbb{E}[\lambda(H_{h_n}^\top(X_i^s)\theta_1)-\lambda(H_{h_n}^\top(X_i^s)\theta_2)].$$

We have

$$\begin{aligned}
&\sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{\sum_{i\in I_1(s)}\xi_i\overline{\Delta}_1(\tau,s,X_i)}{n_1^w(s)}-\frac{\sum_{i\in I_0(s)}\xi_i\overline{\Delta}_1(\tau,s,X_i)}{n_0^w(s)}\right|\\
&\leq \sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{\sum_{i\in I_1(s)}\xi_i[\overline{\Delta}_1(\tau,s,X_i)-H(\theta_{1,s}^{NP}(\tau),\hat{\theta}_{1,s}^{NP}(\tau))-\mathbb{E}(R_1(\tau,s,X_i)|S_i=s)]}{n_1^w(s)}\right|
\end{aligned}$$

$$+\sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{\sum_{i\in I_0(s)}\xi_i[\overline{\Delta}_1(\tau,s,X_i)-H(\theta^{NP}_{1,s}(\tau),\hat{\theta}^{NP}_{1,s}(\tau))-\mathbb{E}(R_1(\tau,s,X_i)|S_i=s)]}{n^w_0(s)}\right| \tag{M.4}$$

We aim to bound the first term on the RHS of (M.4). Note for any $\varepsilon>0$, there exists a constant $M>0$ such that

$$\mathbb{P}\left(\sup_{\tau\in\Upsilon}||\hat{\theta}^{NP}_{a,s}(\tau)-\theta^{NP}_{a,s}(\tau)||_2\leq M\sqrt{h_n\log(n)/n}\right)\geq 1-\varepsilon.$$

On the set $\mathcal{A}(\varepsilon)=\{\sup_{\tau\in\Upsilon}||\hat{\theta}^{NP}_{a,s}(\tau)-\theta^{NP}_{a,s}(\tau)||_2\leq M\sqrt{h_n\log(n)/n}\}$, we have

$$\begin{aligned}
&\sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{\sum_{i\in I_1(s)}\xi_i[\overline{\Delta}_1(\tau,s,X_i)-H(\theta^{NP}_{1,s}(\tau),\hat{\theta}^{NP}_{1,s}(\tau))-\mathbb{E}(R_1(\tau,s,X_i)|S_i=s)]}{n^w_1(s)}\right|\\
&\leq \sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{\sum_{i\in I_1(s)}\xi_i[\lambda(H^\top_{h_n}(X_i)\theta^{NP}_{1,s}(\tau))-\lambda(H^\top_{h_n}(X_i)\hat{\theta}^{NP}_{1,s}(\tau))-H(\theta^{NP}_{1,s}(\tau),\hat{\theta}^{NP}_{1,s}(\tau))]}{n_1(s)}\right|\\
&+\sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{\sum_{i\in I_1(s)}\xi_i\left[R_1(\tau,s,X_i)-\mathbb{E}(R_1(\tau,s,X_i)|S_i=s)\right]}{n^w_1(s)}\right|\\
&\leq\frac{n_1(s)}{n^w_1(s)}\left[\sup_{s\in\mathcal{S},\theta_1,\theta_2\in\Re^{h_n},||\theta_1-\theta_2||_2\leq M\sqrt{h_n\log(n)/n}}\left|\frac{\sum_{i\in I_1(s)}\xi_i[\lambda(H^\top_{h_n}(X_i)\theta_1)-\lambda(H^\top_{h_n}(X_i)\theta_2)-H(\theta_1,\theta_2)]}{n_1(s)}\right|\right.\\
&\left.+\sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{\sum_{i\in I_1(s)}\xi_i\left[R_1(\tau,s,X_i)-\mathbb{E}(R_1(\tau,s,X_i)|S_i=s)\right]}{n_1(s)}\right|\right]\\
&\equiv\frac{n_1(s)}{n^w_1(s)}(D_1+D_2).
\end{aligned}$$

For D_1, we have

$$\begin{aligned}
&D_1|\{A_i,S_i\}_{i\in[n]}\\
&\stackrel{d}{=}\sup\left|\frac{\sum_{i=N(s)}^{N(s)+n_1(s)}\xi^s_i[\lambda(H^\top_{h_n}(X^s_i)\theta_1)-\lambda(H^\top_{h_n}(X^s_i)\theta_2)-H(\theta_1,\theta_2)]}{n_1(s)}\right||\{A_i,S_i\}_{i\in[n]}\\
&\stackrel{d}{=}||\mathbb{P}_{n_1(s)}-\mathbb{P}||_{\mathcal{F}}|\{A_i,S_i\}_{i\in[n]},
\end{aligned}$$

where the supremum in the first equality is taken over $\{s\in\mathcal{S},\theta_1,\theta_2\in\Re^{h_n},||\theta_1-\theta_2||_2\leq M\sqrt{h_n\log(n)/n}\}$ and

$$\mathcal{F}=\left\{\begin{array}{c}\xi^s_i[\lambda(H^\top_{h_n}(X^s_i)\theta_1)-\lambda(H^\top_{h_n}(X^s_i)\theta_2)-H(\theta_1,\theta_2)]:\\ s\in\mathcal{S},\theta_1,\theta_2\in\Re^{h_n},||\theta_1-\theta_2||_2\leq M\sqrt{h_n\log(n)/n}\end{array}\right\}$$

with the envelope $F=2\xi^s_i$. We further note that $||\max_{i\in[n]}2\xi^s_i||_{\mathbb{P},2}\leq C\log(n)$,

$$\sup_{f\in\mathcal{F}}\mathbb{E}f^2\leq\sup\mathbb{E}(H^\top_{h_n}(X^s_i)(\theta_1-\theta_2))^2\leq\kappa_2M^2h_n\log(n)/n,$$

and

$$\sup_Q N(\mathcal{F}, e_Q, \varepsilon||F||_{Q,2}) \leq \left(\frac{a}{\varepsilon}\right)^{ch_n},$$

where a, c are two fixed constants. Therefore, by Chernozhukov et al. (2014, Corollary 5.1),

$$\mathbb{E}\left[||\mathbb{P}_{n_1(s)} - \mathbb{P}||_{\mathcal{F}}|\{A_i, S_i\}_{i\in[n]}\right] = O_p\left(h_n \log(n)/n + \frac{h_n \log^2(n)}{n}\right) = o_p(n^{-1/2}),$$

which implies $D_1 = o_p(n^{-1/2})$.

Similarly, we have

$$\begin{aligned}D_2|\{A_i, S_i\}_{i\in[n]} \stackrel{d}{=} & \sup_{\tau\in\Upsilon, s\in\mathcal{S}} \left|\frac{\sum_{i=N(s)}^{N(s)+n_1(s)} \xi_i^s \left[R_1(\tau, s, X_i^s) - \mathbb{E}(R_1(\tau, s, X_i^s))\right]}{n_1(s)}\right| |\{A_i, S_i\}_{i\in[n]} \\ = & ||\mathbb{P}_{n_1(s)} - \mathbb{P}||_{\mathcal{F}}|\{A_i, S_i\}_{i\in[n]}|\{A_i, S_i\}_{i\in[n]},\end{aligned}$$

where $\mathcal{F} = \{\xi_i^s[\tau - m_1(\tau, s, X_i^s) - \lambda(H_{h_n}^\top(X_i^s)\theta_{1,s}^{NP}(\tau))] : \tau \in \Upsilon\}$ with an envelope $F = \xi_i^s$. In addition, we note $\mathcal{F}$ is nested in

$$\widetilde{\mathcal{F}} = \{\xi_i^s[\tau - m_1(\tau, s, X_i^s) - \lambda(H_{h_n}^\top(X_i^s)\theta_1)] : \tau \in \Upsilon, \theta_1 \in \Re^{h_n}\},$$

so that

$$\sup_Q N(\mathcal{F}, e_Q, \varepsilon||F||_{Q,2}) \leq \sup_Q N(\widetilde{\mathcal{F}}, e_Q, \varepsilon||F||_{Q,2}) \leq \left(\frac{a}{\varepsilon}\right)^{ch_n}.$$

Last,

$$\sup_{f\in\mathcal{F}} \mathbb{E}f^2 = \sup_{\tau\in\Upsilon, a=0,1, s\in\mathcal{S}} \mathbb{E}R_1^2(\tau, s, X_i^s) = O(h_n \log(n)/n).$$

by Chernozhukov et al. (2014, Corollary 5.1),

$$\mathbb{E}\left[||\mathbb{P}_{n_1(s)} - \mathbb{P}||_{\mathcal{F}}|\{A_i, S_i\}_{i\in[n]}\right] = O_p\left(h_n \log(n)/n + \frac{h_n \log^2(n)}{n}\right) = o_p(n^{-1/2}),$$

which implies $D_2 = o_p(n^{-1/2})$. This leads to (M.4).

Step 3. Note $|m_a(\tau_1, s, X_i)| \leq 1$ and

$$\begin{aligned}& |m_a(\tau_1, s, X_i) - m_a(\tau_2, s, X_i)| \\ \leq & |\tau_1 - \tau_2| + |\mathbb{P}(Y_i(a) \leq q_a(\tau_1)|X_i, S_i = s) - \mathbb{P}(Y_i(a) \leq q_a(\tau_2)|X_i, S_i = s)| \\ \leq & \left(1 + \frac{\sup_y f_a(y|X_i, S_i = s)}{\inf_{\tau\in\Upsilon} f_a(q_a(\tau))}\right)|\tau_1 - \tau_2|\end{aligned}$$

$$\leq C|\tau_1 - \tau_2|.$$

This implies Assumptions 3(ii) and 3(iii).

N Proof of Theorem 5.5

We focus on the case with $a = 1$. Note

$$\{X_i, Y_i(1)\}_{i \in I_1(s)} | \{A_i, S_i\}_{i \in [n]} \stackrel{d}{=} \{X_i^s, Y_i^s(1)\}_{i=N(s)+1}^{N(s)+n_1(s)} | \{A_i, S_i\}_{i \in [n]}.$$

where $\{X_i^s, Y_i(1)^s\}_{i \in [n]}$ is an i.i.d. sequence that is independent of $\{A_i, S_i\}_{i \in [n]}$. Therefore,

$$\hat{\theta}_{1,s}^{HD}(q) | \{A_i, S_i\}_{i \in [n]} \stackrel{d}{=} \arg\min_{\theta_a} \frac{-1}{n_1(s)} \sum_{i=N(s)+1}^{N(s)+n_1(s)} \Bigg[1\{Y_i^s(1) \leq q\} \log(\lambda(H_{p_n}(X_i)^\top \theta_a)) \\ + 1\{Y_i^s(1) > q\} \log(1 - \lambda(H_{p_n}^\top (X_i^s)\theta_a)) \Bigg] + \frac{\varrho_{n,1}(s)}{n_1(s)} ||\hat{\Omega}\theta_a||_1 \Bigg| \{A_i, S_i\}_{i \in [n]},$$

and Assumption 12(vi) implies

$$0 < \kappa_1 \leq \inf_{a=0,1, s \in \mathcal{S}, |v|_0 \leq h_n \ell_n} \frac{v^T \left(\frac{1}{n_1(s)} \sum_{i=N(s)+1}^{N(s)+n_1(s)} H_{p_n}(X_i^s) H_{p_n}^\top (X_i^s) \right) v}{||v||_2^2} \\ \leq \sup_{a=0,1, s \in \mathcal{S}, |v|_0 \leq h_n \ell_n} \frac{v^T \left(\frac{1}{n_1(s)} \sum_{i=N(s)+1}^{N(s)+n_1(s)} H_{p_n}(X_i^s) H_{p_n}^\top (X_i^s) \right) v}{||v||_2^2} \leq \kappa_2 < \infty,$$

and

$$0 < \kappa_1 \leq \inf_{a=0,1, s \in \mathcal{S}, |v|_0 \leq h_n \ell_n} \frac{v^T \mathbb{E}(H_{p_n}(X_i^s) H_{p_n}^\top (X_i^s)) v}{||v||_2^2} \\ \leq \sup_{a=0,1, s \in \mathcal{S}, |v|_0 \leq h_n \ell_n} \frac{v^T \mathbb{E}(H_{p_n}(X_i^s) H_{p_n}^\top (X_i^s)) v}{||v||_2^2} \leq \kappa_2 < \infty.$$

In addition, we have $n_1(s)/n \stackrel{a.s.}{\to} \pi(s)p(s) > 0$. Therefore, based on the results established by Belloni et al. (2017), we have, conditionally on $\{A_i, S_i\}_{i \in [n]}$, and thus, unconditionally,

$$\sup_{a=0,1, q \in \mathcal{Q}_a^\varepsilon, s \in \mathcal{S}} ||\hat{\theta}_{a,s}^{HD}(q) - \theta_{a,s}^{HD}(q)||_2 = O_p\left(\sqrt{\frac{h_n \log(p_n)}{n}} \right),$$

$$\sup_{a=0,1, q \in \mathcal{Q}_a^\varepsilon, s \in \mathcal{S}} ||\hat{\theta}_{a,s}^{post}(q) - \theta_{a,s}^{HD}(q)||_2 = O_p\left(\sqrt{\frac{h_n \log(p_n)}{n}} \right),$$

$$\sup_{a=0,1,q\in\mathcal{Q}^{\varepsilon}_a,s\in\mathcal{S}}||\hat{\theta}^{HD}_{a,s}(q)|| = O_p(h_n),$$

and

$$\sup_{a=0,1,q\in\mathcal{Q}^{\varepsilon}_a,s\in\mathcal{S}}||\hat{\theta}^{post}_{a,s}(q)|| = O_p(h_n).$$

In the following, we prove the results when $\hat{\theta}^{HD}_{a,s}(q)$ is used. The results corresponding to $\hat{\theta}^{post}_{a,s}(q)$ can be proved in the same manner and are therefore omitted. Recall

$$\begin{aligned}
&\overline{\Delta}_1(\tau,s,X_i) = \hat{m}_1(\tau,s,X_i) - \overline{m}_1(\tau,s,X_i)\\
&= \mathbb{P}(Y_i(1)\le q_1(\tau)|X_i,S_i=s) - \lambda(H_{p_n}(X_i)^\top\hat{\theta}^{HD}_{1,s}(\hat{q}_1(\tau)))\\
&= [\mathcal{M}_1(q_1(\tau),s,X_i) - \mathcal{M}_1(\hat{q}_1(\tau),s,X_i) + r_a(\hat{q}_1(\tau),s,X_i)]\\
&+ \left[\lambda(H_{p_n}(X_i)^\top\theta^{HD}_{1,s}(\hat{q}_1(\tau))) - \lambda(H_{p_n}(X_i)^\top\hat{\theta}^{HD}_{1,s}(\hat{q}_1(\tau)))\right]\\
&\equiv \mathcal{R}_{a,s}(q_1(\tau),\hat{q}_1(\tau),X_i) + \lambda(H_{p_n}(X_i)^\top\theta^{HD}_{1,s}(\hat{q}_1(\tau))) - \lambda(H_{p_n}(X_i)^\top\hat{\theta}^{HD}_{1,s}(\hat{q}_1(\tau))),
\end{aligned}$$

where

$$\mathcal{R}_{a,s}(q,q',X_i) = \mathcal{M}_1(q,s,X_i) - \mathcal{M}_1(q',s,X_i) + r_a(q',s,X_i).$$

Let

$$H_\lambda(\theta_1,\theta_2,s) = \mathbb{E}[\lambda(H_{p_n}(X_i)^\top\theta_1) - \lambda(H_{p_n}(X_i)^\top\theta_2)|S_i=s],$$

and

$$H_R(q,q',s) = \mathbb{E}(\mathcal{R}_{a,s}(q,q',X_i)|S_i=s).$$

Then, we have

$$\begin{aligned}
&\sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{\sum_{i\in I_1(s)}\xi_i\overline{\Delta}_1(\tau,s,X_i)}{n^w_1(s)} - \frac{\sum_{i\in I_0(s)}\xi_i\overline{\Delta}_1(\tau,s,X_i)}{n^w_0(s)}\right|\\
&\le \sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{\sum_{i\in I_1(s)}\xi_i[\overline{\Delta}_1(\tau,s,X_i) - H_\lambda(\hat{\theta}^{HD}_{1,s}(\hat{q}_1(\tau)),\theta^{HD}_{1,s}(\hat{q}_1(\tau)),s) - H_R(q_1(\tau),\hat{q}_1(\tau),s)]}{n^w_1(s)}\right|\\
&+ \sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{\sum_{i\in I_0(s)}\xi_i[\overline{\Delta}_1(\tau,s,X_i) - H_\lambda(\hat{\theta}^{HD}_{1,s}(\hat{q}_1(\tau)),\theta^{HD}_{1,s}(\hat{q}_1(\tau)),s) - H_R(q_1(\tau),\hat{q}_1(\tau),s)]}{n^w_0(s)}\right| \quad \text{(N.1)}
\end{aligned}$$

We aim to bound the first term on the RHS of (N.1). Note for any $\varepsilon > 0$, there exists a constant

$M > 0$ such that

$$\mathbb{P}\left(\begin{array}{c}\sup_{q\in\mathcal{Q}_1^\varepsilon}||\hat{\theta}_{1,s}^{HD}(q)-\theta_{1,s}^{HD}(q)||_2 \le M\sqrt{\frac{h_n\log(p_n)}{n}},\ \sup_{q\in\mathcal{Q}_1^\varepsilon}||\hat{\theta}_{1,s}^{HD}(q)||_0 \le Mh_n,\\ \sup_{\tau\in\Upsilon}|\hat{q}_1(\tau)-q_1(\tau)| \le Mn^{-1/2}\end{array}\right) \ge 1-\varepsilon.$$

On the set

$$\mathcal{A}(\varepsilon) = \left\{\begin{array}{c}\sup_{q\in\mathcal{Q}_1^\varepsilon}||\hat{\theta}_{1,s}^{HD}(q)-\theta_{1,s}^{HD}(q)||_2 \le M\sqrt{\frac{h_n\log(p_n)}{n}},\ \sup_{q\in\mathcal{Q}_1^\varepsilon}||\hat{\theta}_{1,s}^{HD}(q)||_0 \le Mh_n,\\ \sup_{\tau\in\Upsilon}|\hat{q}_1(\tau)-q_1(\tau)| \le Mn^{-1/2}\end{array}\right\},$$

we have

$$\begin{aligned}
&\sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{\sum_{i\in I_1(s)}\xi_i[\overline{\Delta}_1(\tau,s,X_i)-H_\lambda(\hat{\theta}_{1,s}^{HD}(\hat{q}_1(\tau)),\theta_{1,s}^{HD}(\hat{q}_1(\tau)),s)-H_R(q_1(\tau),\hat{q}_1(\tau),s)]}{n_1^w(s)}\right|\\
&\le \sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{\sum_{i\in I_1(s)}\xi_i[\lambda(H_{h_n}^\top(X_i)\theta_{1,s}^{NP}(\tau))-\lambda(H_{h_n}^\top(X_i)\hat{\theta}_{1,s}^{NP}(\tau))-H_\lambda(\theta_{1,s}^{NP}(\tau),\hat{\theta}_{1,s}^{NP}(\tau),s)]}{n_1(s)}\right|\\
&+\sup_{\tau\in\Upsilon,s\in\mathcal{S}}\left|\frac{\sum_{i\in I_1(s)}\xi_i\left[\mathcal{R}_{1,s}(q_1(\tau),\hat{q}_1(\tau),X_i)-\mathbb{E}(\mathcal{R}_{1,s}(q_1(\tau),\hat{q}_1(\tau),X_i)|S_i=s)\right]}{n_1^w(s)}\right|\\
&\le \frac{n_1(s)}{n_1^w(s)}\left[\sup\left|\frac{\sum_{i\in I_1(s)}\xi_i[\lambda(H_{h_n}^\top(X_i)\theta_1)-\lambda(H_{h_n}^\top(X_i)\theta_2)-H_\lambda(\theta_1,\theta_2,s)]}{n_1(s)}\right|\right.\\
&\left.+\sup_{q,q'\in\mathcal{Q}_1^\varepsilon,|q-q'|\le Mn^{-1/2},s\in\mathcal{S}}\left|\frac{\sum_{i\in I_1(s)}\xi_i\left[\mathcal{R}_{1,s}(q,q',X_i)-\mathbb{E}(\mathcal{R}_{1,s}(q,q',X_i)|S_i=s)\right]}{n_1(s)}\right|\right]\\
&\equiv \frac{n_1(s)}{n_1^w(s)}(D_1+D_2),
\end{aligned}$$

where the first supremum in the second inequality is taken over $\{s\in\mathcal{S},\theta_1,\theta_2\in\Re^{h_n},||\theta_1-\theta_2||_2\le M\sqrt{h_n\log(n)/n},||\theta_1||_0+||\theta_2||_0\le Mh_n\}$. Denote

$$\mathcal{F} = \left\{\begin{array}{c}\xi_i^s[\lambda(H_{h_n}^\top(X_i^s)\theta_1)-\lambda(H_{h_n}^\top(X_i^s)\theta_2)-H_\lambda(\theta_1,\theta_2,s)]:\\ s\in\mathcal{S},\theta_1,\theta_2\in\Re^{h_n},||\theta_1-\theta_2||_2\le M\sqrt{h_n\log(n)/n},||\theta_1||_0+||\theta_2||_0\le Mh_n\end{array}\right\}$$

with the envelope $F=2\xi_i^s$. We further note that $||\max_{i\in[n]}2\xi_i^s||_{\mathbb{P},2}\le C\log(n)$,

$$\sup_{f\in\mathcal{F}}\mathbb{E}f^2 \le \sup\mathbb{E}(H_{h_n}^\top(X_i^s)(\theta_1-\theta_2))^2 \le \kappa_2 M^2 h_n\log(p_n)/n,$$

and

$$\sup_Q N(\mathcal{F},e_Q,\varepsilon||F||_{Q,2}) \le \left(\frac{ap_n}{\varepsilon}\right)^{ch_n},$$

where a, c are two fixed constants. Therefore, Lemma P.2 implies

$$D_1 = O_p\left(\frac{h_n \log(p_n)}{n} + \frac{h_n \log(n)\log(p_n)}{n}\right) = o_p(n^{-1/2}).$$

Similarly, denote

$$\mathcal{F} = \left\{\xi_i^s\left[\mathcal{M}_1(q, s, X_i) - \mathcal{M}_1(q', s, X_i) + r_a(q', s, X_i)\right] : q, q' \in \mathcal{Q}_1^{\varepsilon}, |q - q'| \leq Mn^{-1/2}, s \in \mathcal{S},\right\}$$

with an envelope $F = \xi_i^s$. In addition, note that $\mathcal{F}$ is nested in

$$\widetilde{\mathcal{F}} = \{\xi_i^s[\mathcal{M}_1(q, s, X_i^s) - \lambda(H_{h_n}^\top(X_i^s)\theta_1)] : q \in \mathcal{Q}_1^{\varepsilon}, \theta_1 \in \Re^{p_n}, ||\theta_1||_0 \leq h_n\},$$

with the same envelope. Hence,

$$\sup_Q N(\mathcal{F}, e_Q, \varepsilon||F||_{Q,2}) \leq \sup_Q N(\widetilde{\mathcal{F}}, e_Q, \varepsilon||F||_{Q,2}) \leq \left(\frac{ap_n}{\varepsilon}\right)^{ch_n}.$$

Last,

$$\sup_{f\in\mathcal{F}} \mathbb{E}f^2 \leq C \sup_{q,q'\in\mathcal{Q}_1^{\varepsilon}, |q-q'|\leq Mn^{-1/2}, s\in\mathcal{S}} (|q - q'|^2 + \mathbb{E}r_a^2(q', s, X_i^s)) = O(h_n \log(p_n)/n).$$

Therefore, Lemma P.2 implies

$$D_2 = O_p\left(\frac{h_n \log(p_n)}{n} + \frac{h_n \log(n)\log(p_n)}{n}\right) = o_p(n^{-1/2}).$$

This leads to (N.1). We can establish Assumption 3(i) in the same manner. Assumptions 3(ii) and 3(iii) can be established by the same argument used in Step 3 of the proof of Theorem 5.4. This concludes the proof of Theorem 5.5.

O Proof of Proposition D.1

We divide the proof into two steps. In the first step, we show $d_H(\hat{\Theta}_s^{\varepsilon_n}(\tau), \Theta_s(\tau) \cap \Theta) = o_p(1)$ where $d_H(\cdot, \cdot)$ denotes the Hausdorff distance. In the second step, we show $\hat{\theta}_{a,s}^*(\tau) \xrightarrow{p} \theta_{a,s}^*(\tau)$.

Step 1. For some $\delta_0 > 0$, let

$$\eta = \inf_{(\theta_1,\theta_0)\in\Theta, d_H((\theta_1,\theta_0),\Theta_s(\tau)\cap\Theta)\geq\delta_0} (Q(s, \tau, \theta_1, \theta_0) - \inf_{(\theta_1,\theta_0)\in\Theta_s(\tau)\cap\Theta} Q(s, \tau, \theta_1, \theta_0)).$$

Because for a fixed τ, $\{(\theta_1, \theta_0) \in \Theta, d_H((\theta_1, \theta_0), \Theta_s(\tau)) \geq \delta_0\}$ is compact and $Q(s, \tau, \theta_1, \theta_0)$ is continuous in (θ_1, θ_0), we have $\eta > 0$. On the other hand, for any $(\theta_{1,s}(\tau), \theta_{0,s}(\tau)) \in \Theta_s(\tau) \cap \Theta$,

$$(Q(s, \tau, \theta_1, \theta_0) - Q(s, \tau, \theta_{1,s}(\tau), \theta_{1,s}(\tau)))$$

$$\leq (Q_n(s,\tau,\theta_1,\theta_0) - Q_n(s,\tau,\theta_{1,s}(\tau),\theta_{1,s}(\tau))) + 2\Delta_n,$$

where

$$\Delta_n = \sup_{\tau\in\Upsilon, s\in\mathcal{S}, \theta_1,\theta_0\in\Theta} |Q_n(s,\tau,\theta_1,\theta_0) - Q(s,\tau,\theta_1,\theta_0)|.$$

Taking $\inf_{(\theta_{1,s}(\tau),\theta_{0,s}(\tau))\in\Theta_s(\tau)\cap\Theta}$ on both sides, we have

$$\begin{aligned}
&(Q(s,\tau,\theta_1,\theta_0) - \inf_{(\theta_{1,s}(\tau),\theta_{0,s}(\tau))\in\Theta_s(\tau)\cap\Theta} Q(s,\tau,\theta_{1,s}(\tau),\theta_{1,s}(\tau))) \\
&\leq (Q_n(s,\tau,\theta_1,\theta_0) - \inf_{(\theta_{1,s}(\tau),\theta_{0,s}(\tau))\in\Theta_s(\tau)\cap\Theta} Q(s,\tau,\theta_{1,s}(\tau),\theta_{1,s}(\tau))) + 2\Delta_n,
\end{aligned}$$

Suppose there exist $(\hat{\theta}_{1,s}(\tau),\hat{\theta}_{0,s}(\tau)) \in \hat{\Theta}_s^{\varepsilon_n}(\tau)$ such that $d_H((\hat{\theta}_{1,s}(\tau),\hat{\theta}_{0,s}(\tau),\Theta_s(\tau)\cap\Theta) \geq \delta_0$. Then, because $\Delta_n = o_p(\varepsilon_n)$ as shown in Lemma P.6, we have

$$Q_n(s,\tau,\hat{\theta}_{1,s}(\tau),\hat{\theta}_{0,s}(\tau)) - \inf_{(\theta_{1,s}(\tau),\theta_{0,s}(\tau))\in\Theta_s(\tau)\cap\Theta} Q_n(s,\tau,\theta_{1,s}(\tau),\theta_{1,s}(\tau)) \leq \varepsilon_n,$$

and

$$\begin{aligned}
\eta &\leq (Q(s,\tau,\hat{\theta}_{1,s}(\tau),\hat{\theta}_{0,s}(\tau)) - \inf_{(\theta_{1,s}(\tau),\theta_{0,s}(\tau))\in\Theta_s(\tau)\cap\Theta} Q(s,\tau,\theta_{1,s}(\tau),\theta_{1,s}(\tau))) \\
&\leq Q_n(s,\tau,\hat{\theta}_{1,s}(\tau),\hat{\theta}_{0,s}(\tau)) - \inf_{(\theta_{1,s}(\tau),\theta_{0,s}(\tau))\in\Theta_s(\tau)\cap\Theta} Q_n(s,\tau,\theta_{1,s}(\tau),\theta_{1,s}(\tau)) + 2\Delta_n \leq 2\Delta_n + \varepsilon_n.
\end{aligned}$$

Therefore,

$$\mathbb{P}\left(\exists(\hat{\theta}_{1,s}(\tau),\hat{\theta}_{0,s}(\tau)) \in \hat{\Theta}_s^{\varepsilon_n}(\tau), d_H((\hat{\theta}_{1,s}(\tau),\hat{\theta}_{0,s}(\tau),\Theta_s(\tau)\cap\Theta) \geq \delta_0)\right) \to 0,$$

or equivalently,

$$\sup_{(\hat{\theta}_{1,s}(\tau),\hat{\theta}_{0,s}(\tau))\in\hat{\Theta}_s^{\varepsilon_n}(\tau)} \inf_{(\theta_{1,s}(\tau),\theta_{0,s}(\tau))\in\Theta_s(\tau)\cap\Theta} \sqrt{||\hat{\theta}_{1,s}(\tau)-\theta_{1,s}(\tau)||_2^2 + ||\hat{\theta}_{0,s}(\tau)-\theta_{0,s}(\tau)||_2^2} \xrightarrow{p} 0.$$

Next, note that for any $(\theta_{1,s}(\tau),\theta_{0,s}(\tau)) \in \Theta_s(\tau)\cap\Theta$ and $(\theta'_{1,s}(\tau),\theta'_{0,s}(\tau)) \in \Theta$, we have

$$\begin{aligned}
&Q_n(s,\tau,\theta_{1,s}(\tau),\theta_{0,s}(\tau)) - Q_n(s,\tau,\theta'_{1,s}(\tau),\theta'_{0,s}(\tau)) \\
&\leq Q(s,\tau,\theta_{1,s}(\tau),\theta_{0,s}(\tau)) - Q(s,\tau,\theta'_{1,s}(\tau),\theta'_{0,s}(\tau)) + 2\Delta_n \leq 2\Delta_n = o_p(\varepsilon_n),
\end{aligned}$$

where the last equality is due to the second part of Lemma P.6. Therefore, by the definition of $\hat{\Theta}_s^{\varepsilon_n}(\tau)$,

$$\mathbb{P}(\Theta_s(\tau)\cap\Theta \subset \hat{\Theta}_s^{\varepsilon_n}(\tau)) \to 1,$$

which implies

$$\sup_{(\theta_{1,s}(\tau),\theta_{0,s}(\tau))\in\Theta_s(\tau)\cap\Theta}\ \inf_{(\hat{\theta}_{1,s}(\tau),\hat{\theta}_{0,s}(\tau))\in\hat{\Theta}_s^{\varepsilon_n}(\tau)} \sqrt{||\hat{\theta}_{1,s}(\tau)-\theta_{1,s}(\tau)||_2^2+||\hat{\theta}_{0,s}(\tau)-\theta_{0,s}(\tau)||_2^2} \xrightarrow{p} 0.$$

This concludes the first step of the proof.

Step 2. Because $\Theta_s(\tau)\cap\Theta \subset \hat{\Theta}_s^{\varepsilon_n}(\tau)$ with probability approaching one, we have

$$\sqrt{||\hat{\theta}_{1,s}^*(\tau)||_2^2+||\hat{\theta}_{0,s}^*(\tau)||_2^2} \le \sqrt{||\theta_{1,s}^*(\tau)||_2^2+||\theta_{0,s}^*(\tau)||_2^2} \ \ w.p.a.1.$$

On the other hand, for $\hat{\theta}_{a,s}^*(\tau)$, we can find $(\tilde{\theta}_{1,s}^*(\tau),\tilde{\theta}_{0,s}^*(\tau))\in\Theta_s(\tau)\cap\Theta$ such that

$$\sqrt{||\hat{\theta}_{1,s}^*(\tau)-\tilde{\theta}_{1,s}^*(\tau)||_2^2+||\hat{\theta}_{0,s}^*(\tau)-\tilde{\theta}_{0,s}^*(\tau)||_2^2} \le d_H(\hat{\Theta}_s^{\varepsilon_n}(\tau),\Theta_s(\tau)\cap\Theta)=o_p(1), \tag{O.1}$$

and thus,

$$\begin{aligned}\sqrt{||\hat{\theta}_{1,s}^*(\tau)||_2^2+||\hat{\theta}_{0,s}^*(\tau)||_2^2} \ge & \sqrt{||\tilde{\theta}_{1,s}^*(\tau)||_2^2+||\tilde{\theta}_{0,s}^*(\tau)||_2^2}-d_H(\hat{\Theta}_s^{\varepsilon_n}(\tau),\Theta_s(\tau)\cap\Theta)\\ = & \sqrt{||\theta_{1,s}^*(\tau)||_2^2+||\theta_{0,s}^*(\tau)||_2^2}-o_p(1).\end{aligned}$$

Therefore,

$$\sqrt{||\hat{\theta}_{1,s}^*(\tau)||_2^2+||\hat{\theta}_{0,s}^*(\tau)||_2^2}-\sqrt{||\theta_{1,s}^*(\tau)||_2^2+||\theta_{0,s}^*(\tau)||_2^2}=o_p(1),$$

and

$$\sqrt{||\tilde{\theta}_{1,s}^*(\tau)||_2^2+||\tilde{\theta}_{0,s}^*(\tau)||_2^2} \xrightarrow{p} \sqrt{||\theta_{1,s}^*(\tau)||_2^2+||\theta_{0,s}^*(\tau)||_2^2}. \tag{O.2}$$

Last, note that $\Theta_s(\tau)\cap\Theta$ is compact because $Q(s,\tau,\theta_1,\theta_2)$ is continuous in (θ_1,θ_2). We also note that the Euclidean distance $d(\theta_1,\theta_0)=\sqrt{||\theta_1||_2^2+||\theta_0||_2^2}$ is a continuous function and $(\theta_{1,s}^*(\tau),\theta_{0,s}^*(\tau))$ is uniquely defined over $\Theta_s(\tau)\cap\Theta$ by Assumption D.1. Then, for any $\delta_0>0$,

$$\eta=\inf_{(\theta_1,\theta_0)\in\Theta_s(\tau)\cap\Theta,d(\theta_1-\theta_{1,s}^*(\tau),\theta_0-\theta_{0,s}^*(\tau))\ge\delta_0} d(\theta_1,\theta_0)-d(\theta_{1,s}^*(\tau),\theta_{0,s}^*(\tau))>0,$$

and

$$\mathbb{P}(d(\tilde{\theta}_{1,s}^*(\tau)-\theta_{1,s}^*(\tau),\tilde{\theta}_{0,s}^*(\tau)-\theta_{0,s}^*(\tau))\ge\delta_0)\le\mathbb{P}\left(d(\tilde{\theta}_{1,s}^*(\tau),\tilde{\theta}_{0,s}^*(\tau))-d(\theta_{1,s}^*(\tau),\theta_{0,s}^*(\tau))\ge\eta\right)\to 0,$$

where the last step is due to (O.2). Therefore, we have

$$(\tilde{\theta}_{1,s}^*(\tau),\tilde{\theta}_{0,s}^*(\tau)) \xrightarrow{p} \theta_{1,s}^*(\tau),\theta_{0,s}^*(\tau),$$

which, along with (O.1), further implies

$$(\hat{\theta}^*_{1,s}(\tau), \hat{\theta}^*_{0,s}(\tau)) \xrightarrow{p} \theta^*_{1,s}(\tau), \theta^*_{0,s}(\tau)$$

P Technical Lemmas

The first lemma was established in Zhang and Zheng (2020).

Lemma P.1. *Let S_k be the k-th partial sum of Banach space valued independent identically distributed random variables, then*

$$\mathbb{P}(\max_{1\leq k\leq n} ||S_k|| \geq \varepsilon) \leq 3 \max_{1\leq k\leq n} \mathbb{P}(||S_k|| \geq \varepsilon/3) \leq 9\mathbb{P}(||S_n|| \geq \varepsilon/30).$$

Proof. The first inequality is due to Zhang and Zheng (2020, Lemma E.1) and the second inequality is due to Montgomery-Smith (1993, Theorem 1). □

The next lemma is due to Chernozhukov et al. (2014) with our modification of their maximal inequality to the case with covariate-adaptive randomization.

Lemma P.2. *Suppose Assumption 1 holds. Let $w_i = 1$ or ξ_i defined in Assumption 4. Denote $\mathcal{F}$ as a class of functions of the form $f(x, y_1, y_0)$ where, $f(x, y_1, y_0)$ is a measurable function and $\mathbb{E}(f(X_i, Y_i(1), Y_i(0))|S_i = s) = 0$. Further suppose $\max_{s\in\mathcal{S}} \mathbb{E}(|F_i|^q|S_i = s) < \infty$ for some $q \geq 2$, where*

$$F_i = \sup_{f\in\mathcal{F}} |w_i f(X_i, Y_i(1), Y_i(0))|,$$

$\mathcal{F}$ is of the VC-type with coefficients $(\alpha_n, v_n) > 0$, and $\sup_{f\in\mathcal{F}} \mathbb{E}(f^2|S = s) \leq \sigma_n^2$. Then,

$$\sup_{f\in\mathcal{F}, s\in\mathcal{S}} \frac{1}{\sqrt{n}} \left| \sum_{i\in[n]} A_i 1\{S_i = s\} w_i f(X_i, Y_i(1), Y_i(0)) \right|$$
$$= O_p\left(\sqrt{v_n \sigma_n^2 \log\left(\frac{\alpha_n ||F||_{\mathbb{P},2}}{\sigma}\right)} + \frac{v_n || \max_{i\in[n]} F_i||_{\mathbb{P},2} \log\left(\frac{\alpha_n||F||_{\mathbb{P},2}}{\sigma}\right)}{\sqrt{n}} \right),$$

and

$$\sup_{f\in\mathcal{F}, s\in\mathcal{S}} \frac{1}{\sqrt{n}} \left| \sum_{i\in[n]} (1 - A_i) 1\{S_i = s\} w_i f(X_i, Y_i(1), Y_i(0)) \right|$$
$$= O_p\left(\sqrt{v_n \sigma_n^2 \log\left(\frac{\alpha_n ||F||_{\mathbb{P},2}}{\sigma}\right)} + \frac{v_n || \max_{i\in[n]} F_i||_{\mathbb{P},2} \log\left(\frac{\alpha_n||F||_{\mathbb{P},2}}{\sigma}\right)}{\sqrt{n}} \right).$$

Proof. We focus on establishing the first statement. The proof of the second statement is similar and is omitted. Following Bugni, Canay, and Shaikh (2018), we define the sequence of i.i.d. random variables $\{(w_i^s, X_i^s, Y_i^s(1), Y_i^s(0)) : 1 \leq i \leq n\}$ with marginal distributions equal to the distribution of $(w_i, X_i, Y_i(1), Y_i(0))|S_i = s$. The distribution of $\sum_{i\in[n]} A_i 1\{S_i = s\} w_i f(X_i, Y_i(1), Y_i(0))$ is the same as the counterpart with units ordered by strata and then ordered by $A_i = 1$ first and $A_i = 0$ second within each stratum, i.e.,

$$\begin{aligned}\sum_{i\in[n]} A_i 1\{S_i = s\} w_i f(X_i, Y_i(1), Y_i(0)) &\stackrel{d}{=} \sum_{i=N(s)+1}^{N(s)+n_1(s)} w_i^s f(X_i^s, Y_i^s(1), Y_i^s(0)) \\ &\equiv \Gamma_n^s(N(s)+n_1(s), f) - \Gamma_n^s(N(s)+1, f),\end{aligned}$$

where $N(s) = \sum_{i\in[n]} 1\{S_i < s\}$ and

$$\Gamma_n^s(k, f) = \sum_{i\in[k]} w_i^s f(X_i^s, Y_i^s(1), Y_i^s(0)).$$

Let $\mu_n = \sqrt{v_n \sigma_n^2 \log\left(\frac{\alpha_n ||F||_{\mathbb{P},2}}{\sigma}\right)} + \frac{v_n || \max_{i\in[n]} F_i||_{\mathbb{P},2} \log\left(\frac{\alpha_n ||F||_{\mathbb{P},2}}{\sigma}\right)}{\sqrt{n}}$. Then, for some constant $C > 0$, we have

$$\begin{aligned}&\mathbb{P}\left(\sup_{f\in\mathcal{F}, s\in\mathcal{S}} \frac{1}{\sqrt{n}} \left|\sum_{i\in[n]} A_i 1\{S_i = s\} w_i f(X_i, Y_i(1), Y_i(0))\right| \geq t\mu_n\right) \\ &= \mathbb{P}\left(\sup_{f\in\mathcal{F}, s\in\mathcal{S}} \frac{1}{\sqrt{n}} \left|\Gamma_n^s(N(s)+n_1(s), f) - \Gamma_n^s(N(s)+1, f)\right| \geq t\mu_n\right) \\ &\leq \sum_{s\in\mathcal{S}} \mathbb{P}\left(\max_{1\leq k\leq n} \sup_{f\in\mathcal{F}} \left|\frac{1}{\sqrt{n}} \Gamma_n^s(k, f)\right| \geq t\mu_n/2\right) \\ &\leq \sum_{s\in\mathcal{S}} 9\, \mathbb{P}\left(\sup_{f\in\mathcal{F}} \left|\frac{1}{\sqrt{n}} \Gamma_n^s(n, f)\right| \geq t\mu_n/60\right) \\ &\leq \sum_{s\in\mathcal{S}} \frac{540\, \mathbb{E}\left(\sup_{f\in\mathcal{F}} \left|\frac{1}{\sqrt{n}} \Gamma_n^s(n, f)\right|\right)}{t\mu_n} \\ &= \sum_{s\in\mathcal{S}} \frac{540 \mathbb{E}\left(\sqrt{n} ||\mathbb{P}_n^s - \mathbb{P}^s||_{\mathcal{F}}\right)}{t\mu_n} \\ &\leq C/t,\end{aligned}$$

where $\mathbb{P}_n^s$ and $\mathbb{P}^s$ are the empirical process and expectation w.r.t. i.i.d. data $\{w_i^s, X_i^s, Y_i^s(1), Y_i^s(0)\}_{i\in[n]}$, respectively, the second inequality is due to Lemma P.1, the last equality is due to the fact that

$$\mathbb{E} w_i^s f(X_i^s, Y_i^s(1), Y_i^s(0)) = \mathbb{E}\left(w_i f(X_i, Y_i(1), Y_i(0)) | S_i = s\right) = 0,$$

and the last inequality is due to the fact that, by Chernozhukov et al. (2014, Corollary 5.1),

$$\mathbb{E}\left(\sqrt{n} ||\mathbb{P}_n^s - \mathbb{P}^s||_{\mathcal{F}}\right) \leq C\mu_n.$$

Then, for any $\varepsilon > 0$, we can choose $t \geq C/\varepsilon$ so that

$$\mathbb{P}\left(\sup_{f \in \mathcal{F}, s \in \mathcal{S}} \frac{1}{\sqrt{n}} \left|\sum_{i \in [n]} A_i 1\{S_i = s\} w_i f(X_i, Y_i(1), Y_i(0))\right| \geq t\mu_n\right) \leq \varepsilon,$$

which implies the desired result. $\square$

Lemma P.3. *Suppose Assumptions in Theorem 3.1 hold. Denote*

$$\begin{aligned}
W_{n,1}(\tau) =& \sum_{s \in \mathcal{S}} \frac{1}{\sqrt{n}} \sum_{i \in [n]} A_i 1\{S_i = s\} \phi_1(\tau, s, Y_i(1), X_i) \\
& - \sum_{s \in \mathcal{S}} \frac{1}{\sqrt{n}} \sum_{i \in [n]} (1 - A_i) 1\{S_i = s\} \phi_0(\tau, s, Y_i(0), X_i),
\end{aligned}$$

and

$$W_{n,2}(\tau) = \frac{1}{\sqrt{n}} \sum_{i \in [n]} \phi_s(\tau, S_i).$$

Then, uniformly over $\tau \in \Upsilon$,

$$(W_{n,1}(\tau), W_{n,2}(\tau)) \rightsquigarrow (\mathcal{B}_1(\tau), \mathcal{B}_2(\tau)),$$

where $(\mathcal{B}_1(\tau), \mathcal{B}_2(\tau))$ are two independent Gaussian processes with covariance kernels $\Sigma_1(\tau, \tau')$ and $\Sigma_2(\tau, \tau')$, respectively, such that

$$\begin{aligned}
\Sigma_1(\tau, \tau') =& \mathbb{E}\pi(S_i)\phi_1(\tau, S_i, Y_i(1), X_i)\phi_1(\tau', S_i, Y_i(1), X_i) \\
& + \mathbb{E}(1 - \pi(S_i))\phi_0(\tau, S_i, Y_i(0), X_i)\phi_0(\tau', S_i, Y_i(0), X_i)
\end{aligned}$$

and

$$\Sigma_2(\tau, \tau') = \mathbb{E}\phi_s(\tau, S_i)\phi_s(\tau', S_i).$$

Proof. We follow the general argument in the proof of Bugni et al. (2018, Lemma B.2). We divide the proof into two steps. In the first step, we show that

$$(W_{n,1}(\tau), W_{n,2}(\tau)) \stackrel{d}{=} (W^{\star}_{n,1}(\tau), W_{n,2}(\tau)) + o_p(1),$$

where the $o_p(1)$ term holds uniformly over $\tau \in \Upsilon$, $W^{\star}_{n,1}(\tau) \perp\!\!\!\perp W_{n,2}(\tau)$, and, uniformly over $\tau \in \Upsilon$,

$$W^{\star}_{n,1}(\tau) \rightsquigarrow \mathcal{B}_1(\tau).$$

In the second step, we show that

$$W_{n,2}(\tau) \rightsquigarrow \mathcal{B}_2(\tau)$$

uniformly over $\tau \in \Upsilon$.

Step 1. Recall that we define $\{(X_i^s, Y_i^s(1), Y_i^s(0)) : 1 \leq i \leq n\}$ as a sequence of i.i.d. random variables with marginal distributions equal to the distribution of $(X_i, Y_i(1), Y_i(0))|S_i = s$ and $N(s) = \sum_{i\in[n]} 1\{S_i < s\}$. The distribution of $W_{n,1}(\tau)$ is the same as the counterpart with units ordered by strata and then ordered by $A_i = 1$ first and $A_i = 0$ second within each stratum, i.e.,

$$W_{n,1}(\tau)|\{(A_i, S_i)_{i\in[n]}\} \stackrel{d}{=} \widetilde{W}_{n,1}(\tau)|\{(A_i, S_i)_{i\in[n]}\}$$

where

$$\begin{aligned}\widetilde{W}_{n,1}(\tau) \equiv & \sum_{s\in\mathcal{S}} \frac{1}{\sqrt{n}} \sum_{i=N(s)+1}^{N(s)+n_1(s)} \phi_1(\tau, s, Y_i^s(1), X_i^s) \\ & - \sum_{s\in\mathcal{S}} \frac{1}{\sqrt{n}} \sum_{i=N(s)+n_1(s)+1}^{N(s)+n(s)} \phi_0(\tau, s, Y_i^s(0), X_i^s),\end{aligned}$$

with

$$\begin{aligned}\phi_1(\tau, s, Y_i^s(1), X_i^s) = & \frac{\tau - 1\{Y_i^s(1) \leq q_1(\tau)\} - m_1(\tau, s) - (1-\pi(s))\left(\overline{m}_1(\tau, s, X_i^s) - \overline{m}_1(\tau, s)\right)}{\pi(s) f_1(q_1(\tau))} \\ & - \frac{\left(\overline{m}_0(\tau, s, X_i^s) - \overline{m}_0(\tau, s)\right)}{f_0(q_0(\tau))},\end{aligned}$$

and

$$\begin{aligned}\phi_0(\tau, s, Y_i^s(0), X_i^s) = & \frac{\tau - 1\{Y_i^s(0) \leq q_0(\tau)\} - m_0(\tau, s) - \pi(s)\left(\overline{m}_0(\tau, s, X_i^s) - \overline{m}_0(\tau, s)\right)}{(1-\pi(s)) f_0(q_0(\tau))} \\ & - \frac{\left(\overline{m}_1(\tau, s, X_i^s) - \overline{m}_1(\tau, s)\right)}{f_1(q_1(\tau))}.\end{aligned}$$

As $W_{n,2}(\tau)$ is only a function of $\{S_i\}_{i\in[n]}$, we have

$$(W_{n,1}(\tau), W_{n,2}(\tau)) \stackrel{d}{=} (\widetilde{W}_{n,1}(\tau), W_{n,2}(\tau)).$$

Let $F(s) = \mathbb{P}(S_i < s)$, $p(s) = \mathbb{P}(S_i = s)$, and

$$\begin{aligned}W_{n,1}^{\star}(\tau) \equiv & \sum_{s\in\mathcal{S}} \frac{1}{\sqrt{n}} \sum_{i=\lfloor nF(s)\rfloor+1}^{\lfloor n(F(s)+\pi(s)p(s))\rfloor} \phi_1(\tau, s, Y_i^s(1), X_i^s) \\ & - \sum_{s\in\mathcal{S}} \frac{1}{\sqrt{n}} \sum_{i=\lfloor n(F(s)+\pi(s)p(s))\rfloor+1}^{\lfloor n(F(s)+p(s))\rfloor} \phi_0(\tau, s, Y_i^s(0), X_i^s).\end{aligned}$$

Note $W^{\star}_{n,1}(\tau)$ is a function of $(Y_i^s(1), Y_i^s(0), X_i^s)_{i\in[n],s\in\mathcal{S}}$, which is independent of $\{A_i, S_i\}_{i\in[n]}$ by construction. Therefore,

$$W^{\star}_{n,1}(\tau) \perp\!\!\!\perp W_{n,2}(\tau).$$

Note that

$$\frac{N(s)}{n} \xrightarrow{p} F(s), \quad \frac{n_1(s)}{n} \xrightarrow{p} \pi(s)p(s), \quad \text{and} \quad \frac{n(s)}{n} \xrightarrow{p} p(s).$$

Denote $\Gamma_{n,a}(s,t,\tau) = \sum_{i=1}^{\lfloor nt \rfloor} \frac{1}{n}\phi_a(\tau, s, Y_i^s(a), X_i^s)$ for $a = 0, 1$. In order to show that $\sup_{\tau\in\Upsilon} |\widetilde{W}_{n,1}(\tau) - W^{\star}_{n,1}(\tau)| = o_p(1)$ and $W^{\star}_{n,1}(\tau) \rightsquigarrow \mathcal{B}_1(\tau)$, it suffices to show that (1) for $a = 0, 1$ and $s \in \mathcal{S}$, the stochastic process

$$\{\Gamma_{n,a}(s,t,\tau) : t \in (0,1), \tau \in \Upsilon\}$$

is stochastically equicontinuous and (2) $W^{\star}_{n,1}(\tau) \rightsquigarrow \mathcal{B}_1(\tau)$ converges to $B_1(\tau)$ in finite dimension.

Claim (1). We want to bound

$$\sup |\Gamma_{n,a}(s,t_2,\tau_2) - \Gamma_{n,a}(s,t_1,\tau_1)|,$$

where the supremum is taken over $0 < t_1 < t_2 < t_1 + \varepsilon < 1$ and $\tau_1 < \tau_2 < \tau_1 + \varepsilon$ such that $\tau_1, \tau_1 + \varepsilon \in \Upsilon$. Note that,

$$\begin{aligned}
&\sup |\Gamma_{n,a}(s,t_2,\tau_2) - \Gamma_{n,a}(s,t_1,\tau_1)| \\
\le& \sup_{0<t_1<t_2<t_1+\varepsilon<1, \tau\in\Upsilon} |\Gamma_{n,a}(s,t_2,\tau) - \Gamma_{n,a}(s,t_1,\tau)| + \sup_{t\in(0,1),\tau_1,\tau_2\in\Upsilon,\tau_1<\tau_2<\tau_1+\varepsilon} |\Gamma_{n,a}(s,t,\tau_2) - \Gamma_{n,a}(s,t,\tau_1)|.
\end{aligned} \tag{P.1}$$

Then, for an arbitrary $\delta > 0$, by taking $\varepsilon = \delta^4$, we have

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0<t_1<t_2<t_1+\varepsilon<1,\tau\in\Upsilon} |\Gamma_{n,a}(s,t_2,\tau) - \Gamma_{n,a}(s,t_1,\tau)| \ge \delta\right) \\
&= \mathbb{P}\left(\sup_{0<t_1<t_2<t_1+\varepsilon<1,\tau\in\Upsilon} |\sum_{i=\lfloor nt_1 \rfloor+1}^{\lfloor nt_2 \rfloor} \phi_a(\tau, s, Y_i^s(a), X_i^s)| \ge \sqrt{n}\delta\right) \\
&= \mathbb{P}\left(\sup_{0<t\le\varepsilon,\tau\in\Upsilon} |\sum_{i=1}^{\lfloor nt \rfloor} \phi_a(\tau, s, Y_i^s(a), X_i^s)| \ge \sqrt{n}\delta\right) \\
&\le \mathbb{P}\left(\max_{1\le k\le\lfloor n\varepsilon \rfloor} \sup_{\tau\in\Upsilon} |S_k(\tau)| \ge \sqrt{n}\delta\right) \\
&\le \frac{270\, \mathbb{E} \sup_{\tau\in\Upsilon} |\sum_{i=1}^{\lfloor n\varepsilon \rfloor} \phi_a(\tau, s, Y_i^s(a), X_i^s)|}{\sqrt{n}\delta} \\
&\lesssim \frac{\sqrt{n\varepsilon}}{\sqrt{n}\delta} \lesssim \delta,
\end{aligned}$$

where in the first inequality, $S_k(\tau) = \sum_{i=1}^k \phi_a(\tau, s, Y_i^s(a), X_i^s)$ and the second inequality holds due to Lemma P.1. To see the third inequality, denote

$$\mathcal{F} = \{\phi_a(\tau, s, Y_i^s(a), X_i^s) : \tau \in \Upsilon\}$$

with an envelope function F_i such that by Assumption 3, $||F_i||_{\mathbb{P},q} < \infty$. In addition, by Assumption 3 again and the fact that

$$\{\tau - 1\{Y_i^s(a) \leq q_a(\tau)\} - m_a(\tau, s) : \tau \in \Upsilon\}$$

is of the VC-type with fixed coefficients (α, v), so is $\mathcal{F}$. Then, we have

$$J(1, \mathcal{F}) < \infty,$$

where

$$J(\delta, \mathcal{F}) = \sup_Q \int_0^\delta \sqrt{1 + \log N(\mathcal{F}, L_2(Q), \varepsilon ||F||_{Q,2})} d\varepsilon,$$

$N(\mathcal{F}, L_2(Q), \varepsilon||F||_{Q,2})$ is the covering number, and the supremum is taken over all discrete probability measures Q. Therefore, by van der Vaart and Wellner (1996, Theorem 2.14.1)

$$\frac{270 \mathbb{E} \sup_{\tau \in \Upsilon} |\sum_{i=1}^{\lfloor n\varepsilon \rfloor} \phi_a(\tau, s, Y_i^s(a), X_i^s)|}{\sqrt{n}\delta} \lesssim \frac{\sqrt{\lfloor n\varepsilon \rfloor} \left[\mathbb{E}\sqrt{\lfloor n\varepsilon \rfloor} ||\mathbb{P}_{\lfloor n\varepsilon \rfloor} - \mathbb{P}||_{\mathcal{F}}\right]}{\sqrt{n}\delta} \lesssim \frac{\sqrt{\lfloor n\varepsilon \rfloor} J(1, \mathcal{F})}{\sqrt{n}\delta}.$$

For the second term on the RHS of (P.1), by taking $\varepsilon = \delta^4$, we have

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t \in (0,1), \tau_1, \tau_2 \in \Upsilon, \tau_1 < \tau_2 < \tau_1 + \varepsilon} |\Gamma_{n,a}(s, t, \tau_2) - \Gamma_{n,a}(s, t, \tau_1)| \geq \delta\right) \\
&= \mathbb{P}\left(\max_{1 \leq k \leq n} \sup_{\tau_1, \tau_2 \in \Upsilon, \tau_1 < \tau_2 < \tau_1 + \varepsilon} |S_k(\tau_1, \tau_2)| \geq \sqrt{n}\delta\right) \\
&\leq \frac{270 \, \mathbb{E} \sup_{\tau_1, \tau_2 \in \Upsilon, \tau_1 < \tau_2 < \tau_1 + \varepsilon} |\sum_{i=1}^n (\phi_a(\tau_2, s, Y_i^s(a), X_i^s) - \phi_a(\tau_1, s, Y_i^s(a), X_i^s)|}{\sqrt{n}\delta} \\
&\lesssim \delta \sqrt{\log(\frac{C}{\delta^2})},
\end{aligned}$$

where in the first equality, $S_k(\tau_1, \tau_2) = \sum_{i=1}^k (\phi_a(\tau_2, s, Y_i^s(a), X_i^s) - \phi_a(\tau_1, s, Y_i^s(a), X_i^s))$ and the first inequality is due to Lemma P.1. To see the last inequality, denote

$$\mathcal{F} = \{\phi_a(\tau_2, s, Y_i^s(a), X_i^s) - \phi_a(\tau_1, s, Y_i^s(a), X_i^s) : \tau_1, \tau_2 \in \Upsilon, \tau_1 < \tau_2 < \tau_1 + \varepsilon\}$$

with a constant envelope function F_i such that $||F_i||_{\mathbb{P},q} < \infty$. In addition, due to Assumptions 2.2

and 3.3, one can show that

$$\sup_{f \in \mathcal{F}} \mathbb{E} f^2 \leq c\varepsilon \equiv \sigma^2$$

for some constant $c > 0$. Last, due to Assumption 3.2, $\mathcal{F}$ is of the VC-type with fixed coefficients (α, v). Therefore, by Chernozhukov et al. (2014, Corollary 5.1),

$$\frac{270\mathbb{E}\sup_{\tau_1,\tau_2\in\Upsilon,\tau_1<\tau_2<\tau_1+\varepsilon}|\sum_{i=1}^n(\phi_a(\tau_2,s,Y_i^s(a),X_i^s)-\phi_a(\tau_1,s,Y_i^s(a),X_i^s))|}{\sqrt{n}\delta}$$
$$\lesssim \frac{\sqrt{n}\mathbb{E}||\mathbb{P}_n - \mathbb{P}||_{\mathcal{F}}}{\delta} \lesssim \sqrt{\frac{\sigma^2\log(\frac{C}{\sigma})}{\delta^2}} + \frac{C\log(\frac{C}{\sigma})}{\sqrt{n}\delta} \lesssim \delta\sqrt{\log(\frac{C}{\delta^2})},$$

where the last inequality holds by letting n be sufficiently large. Note that $\delta\sqrt{\log(\frac{C}{\delta^2})} \to 0$ as $\delta \to 0$. This concludes the proof of Claim (1).

Claim (2). For a single τ, by the triangular array CLT,

$$W^\star_{n,1}(\tau) \rightsquigarrow N(0, \Sigma_1(\tau,\tau)),$$

where

$$\begin{aligned}
\Sigma_1(\tau,\tau) &= \lim_{n\to\infty}\sum_{s\in\mathcal{S}}\frac{(\lfloor n(F(s)+\pi(s)p(s))\rfloor - \lfloor nF(s)\rfloor)}{n}\mathbb{E}\phi_1^2(\tau,s,Y_i^s(1),X_i^s) \\
&\quad + \lim_{n\to\infty}\sum_{s\in\mathcal{S}}\frac{(\lfloor n(F(s)+p(s))\rfloor - \lfloor n(F(s)+p(s)\pi(s))\rfloor)}{n}\mathbb{E}\phi_0^2(\tau,s,Y_i^s(0),X_i^s) \\
&= \sum_{s\in\mathcal{S}} p(s)\mathbb{E}(\pi(s)\phi_1^2(\tau,S_i,Y_i(1),X_i) + (1-\pi(s))\phi_0^2(\tau,S_i,Y_i(0),X_i)|S_i = s) \\
&= \mathbb{E}\pi(S_i)\phi_1^2(\tau,S_i,Y_i(1),X_i) + \mathbb{E}(1-\pi(S_i))\phi_0^2(\tau,S_i,Y_i(0),X_i).
\end{aligned}$$

Finite dimensional convergence is proved by the Cramér-Wold device. In particular, we can show that the covariance kernel is

$$\begin{aligned}
\Sigma_1(\tau,\tau') =& \mathbb{E}\pi(S_i)\phi_1(\tau,S_i,Y_i(1),X_i)\phi_1(\tau',S_i,Y_i(1),X_i) \\
&+ \mathbb{E}(1-\pi(S_i))\phi_0(\tau,S_i,Y_i(0),X_i)\phi_0(\tau',S_i,Y_i(0),X_i).
\end{aligned}$$

This concludes the proof of Claim (2), and thereby leads to the desired results in Step 1.

Step 2. As $m_a(\tau,S_i) = \tau - \mathbb{P}(Y_i(a) \leq q_a(\tau)|S_i)$ is Lipschitz continuous in τ with a bounded Lipschitz constant, $\{m_a(\tau,S_i) : \tau \in \Upsilon\}$ is of the VC-type with fixed coefficients (α, v) and a constant envelope function. Therefore, $\{\frac{m_1(\tau,S_i)}{f_1(q_1(\tau))} - \frac{m_0(\tau,S_i)}{f_0(q_0(\tau))} : \tau \in \Upsilon\}$ is a Donsker class and we have

$$W_{n,2}(\tau) \rightsquigarrow \mathcal{B}_2(\tau),$$

where $\mathcal{B}_2(\tau)$ is a Gaussian process with covariance kernel

$$\Sigma_2(\tau,\tau') = \mathbb{E}\left(\frac{m_1(\tau,S_i)}{f_1(q_1(\tau))} - \frac{m_0(\tau,S_i)}{f_0(q_0(\tau))}\right)\left(\frac{m_1(\tau',S_i)}{f_1(q_1(\tau'))} - \frac{m_0(\tau',S_i)}{f_0(q_0(\tau'))}\right) \equiv \mathbb{E}\phi_s(\tau,S_i)\phi_s(\tau',S_i).$$

This concludes the proof. □

Lemma P.4. *Suppose Assumptions in Theorem 4.1 hold and recall* $D_n^w(s) = \sum_{i\in[n]}\xi_i(A_i - \pi(S_i))1\{S_i = s\}$. *Then,* $\max_{s\in\mathcal{S}}|(D_n^w(s) - D_n(s))/n(s)| = o_p(1)$ *and* $\max_{s\in\mathcal{S}}|D_n^w(s)/n^w(s)| = o_p(1)$.

Proof. We note that $n^w(s)/n(s) \xrightarrow{p} 1$ and $D_n(s)/n(s) \xrightarrow{p} 0$. Therefore, we only need to show

$$\frac{D_n^w(s) - D_n(s)}{n(s)} = \sum_{i=1}^n \frac{(\xi_i - 1)(A_i - \pi(s))1\{S_i = s\}}{n(s)} \xrightarrow{p} 0.$$

As $n(s) \to \infty$ a.s., given data,

$$\begin{aligned}\frac{1}{n(s)}\sum_{i=1}^n (A_i - \pi(s))^2 1\{S_i = s\} =& \frac{1}{n}\sum_{i=1}^n \left(A_i - \pi(s) - 2\pi(s)(A_i - \pi(s)) + \pi(s) - \pi^2(s)\right) 1\{S_i = s\} \\ =& \frac{D_n(s) - 2\pi(s)D_n(s)}{n(s)} + \pi(s)(1-\pi(s)) \xrightarrow{p} \pi(s)(1-\pi(s)).\end{aligned}$$

Then, by the Lindeberg CLT, conditionally on data,

$$\frac{1}{\sqrt{n(s)}}\sum_{i=1}^n (\xi_i - 1)(A_i - \pi(s))1\{S_i = s\} \rightsquigarrow N(0, \pi(s)(1-\pi(s))) = O_p(1),$$

and thus

$$\frac{D_n^w(s) - D_n(s)}{n(s)} = O_p(n^{-1/2}(s)) = o_p(1).$$

□

Lemma P.5. *Suppose Assumptions in Theorem 4.1 hold. Then, uniformly over* $\tau \in \Upsilon$ *and conditionally on data,*

$$W_{n,1}^w(\tau) + W_{n,2}^w(\tau) \rightsquigarrow \mathcal{B}(\tau),$$

where $\mathcal{B}(\tau)$ *is a Gaussian process with the covariance kernel*

$$\begin{aligned}&\Sigma(\tau,\tau') = \mathbb{E}\pi(S_i)\phi_1(\tau,S_i,Y_i(1),X_i)\phi_1(\tau',S_i,Y_i(1),X_i) \\ &\quad + \mathbb{E}(1-\pi(S_i))\phi_0(\tau,S_i,Y_i(0),X_i)\phi_0(\tau',S_i,Y_i(0),X_i) + \mathbb{E}\phi_s(\tau,S_i)\phi_s(\tau',S_i).\end{aligned}$$

Proof. We divide the proof into two steps. In the first step, we show the conditional stochastic equicontinuity of $W_{n,1}^w(\tau)$ and $W_{n,2}^w(\tau)$. In the second step, we show the finite-dimensional convergence of $W_{n,1}^w(\tau) + W_{n,2}^w(\tau)$ conditional on data.

Step 1. Following the same idea in the proof of Lemma P.3, we define $\{(\xi_i^s, X_i^s, Y_i^s(1), Y_i^s(0)) : 1 \leq i \leq n\}$ as a sequence of i.i.d. random variables with marginal distributions equal to the distribution of $(\xi_i, X_i, Y_i(1), Y_i(0))|S_i = s$ and $N(s) = \sum_{i\in[n]} 1\{S_i < s\}$. The distribution of $W_{n,1}(\tau)$ is the same as the counterpart with units ordered by strata and then ordered by $A_i = 1$ first and $A_i = 0$ second within each stratum, i.e.,

$$W_{n,1}^w(\tau)|\{(A_i, S_i)_{i\in[n]}\} \stackrel{d}{=} \widetilde{W}_{n,1}^w(\tau)|\{(A_i, S_i)_{i\in[n]}\},$$

and thus,

$$W_{n,1}^w(\tau) \stackrel{d}{=} \widetilde{W}_{n,1}^w(\tau), \tag{P.2}$$

where

$$\begin{aligned}\widetilde{W}_{n,1}^w(\tau) \equiv &\sum_{s\in\mathcal{S}} \frac{1}{\sqrt{n}} \sum_{i=N(s)+1}^{N(s)+n_1(s)} (\xi_i^s - 1)\phi_1(\tau, s, Y_i^s(1), X_i^s) \\ &- \sum_{s\in\mathcal{S}} \frac{1}{\sqrt{n}} \sum_{i=N(s)+n_1(s)+1}^{N(s)+n(s)} (\xi_i^s - 1)\phi_0(\tau, s, Y_i^s(0), X_i^s).\end{aligned}$$

In addition, let

$$\begin{aligned}W_{n,1}^{w\star}(\tau) \equiv &\sum_{s\in\mathcal{S}} \frac{1}{\sqrt{n}} \sum_{i=\lfloor nF(s)\rfloor+1}^{\lfloor n(F(s)+\pi(s)p(s))\rfloor} (\xi_i^s - 1)\phi_1(\tau, s, Y_i^s(1), X_i^s) \\ &- \sum_{s\in\mathcal{S}} \frac{1}{\sqrt{n}} \sum_{i=\lfloor n(F(s)+\pi(s)p(s))\rfloor+1}^{\lfloor n(F(s)+p(s))\rfloor} (\xi_i^s - 1)\phi_0(\tau, s, Y_i^s(0), X_i^s).\end{aligned}$$

Following exactly the same argument as in the proof of Lemma P.3, we have

$$\sup_{\tau\in\Upsilon} |\widetilde{W}_{n,1}^w(\tau) - W_{n,1}^{w\star}(\tau)| = o_p(1). \tag{P.3}$$

and $W_{n,1}^{w\star}(\tau)$ is *unconditionally* stochastically equicontinuous, i.e., for any $\varepsilon > 0$, as $n \to \infty$ followed by $\delta \to 0$, we have

$$\begin{aligned}&\mathbb{E}\mathbb{P}_\xi\left(\sup_{\tau_1,\tau_2\in\Upsilon,\tau_1<\tau_2<\tau_1+\delta} |W_{n,1}^{w\star}(\tau_1) - W_{n,1}^{w\star}(\tau_2)| \geq \varepsilon\right) \\ =&\mathbb{P}\left(\sup_{\tau_1,\tau_2\in\Upsilon,\tau_1<\tau_2<\tau_1+\delta} |W_{n,1}^{w\star}(\tau_1) - W_{n,1}^{w\star}(\tau_2)| \geq \varepsilon\right) \to 0,\end{aligned}$$

where $\mathbb{P}_\xi$ means the probability operator is with respect to the bootstrap weights $\{\xi_i\}_{i\in[n]}$ and is conditional on data. This implies the *unconditional* stochastic equicontinuity of $W_{n,1}^w(\tau)$ due to (P.2) and (P.3), which further implies the *conditional* stochastic equicontinuity of $W_{n,1}^w(\tau)$, i.e., for any $\varepsilon > 0$, as $n \to \infty$ followed by $\delta \to 0$,

$$\mathbb{P}_\xi\left(\sup_{\tau_1,\tau_2\in\Upsilon,\tau_1<\tau_2<\tau_1+\delta}|W^{w\star}_{n,1}(\tau_1) - W^{w\star}_{n,1}(\tau_2)| \geq \varepsilon\right) \xrightarrow{p} 0.$$

By a similar but simpler argument, the *conditional* stochastic equicontinuity of $W^w_{n,2}(\tau)$ holds as well. This concludes the first step.

Step 2. We first show the asymptotic normality of $W^w_{n,1}(\tau) + W^w_{n,2}(\tau)$ conditionally on data for a fixed τ. Note

$$\begin{aligned}
&W^w_{n,1}(\tau) + W^w_{n,2}(\tau)\\
=&\frac{1}{\sqrt{n}}\sum_{i\in[n]}(\xi_i - 1)\left[A_i 1\{S_i = s\}\phi_1(\tau, S_i, Y_i(1), X_i) - (1 - A_i)1\{S_i = s\}\phi_1(\tau, S_i, Y_i(1), X_i) + \phi_s(\tau, S_i)\right]\\
\equiv&\frac{1}{\sqrt{n}}\sum_{i\in[n]}(\xi_i - 1)\mathcal{J}_i(s, \tau).
\end{aligned}$$

Conditionally on data, $\{(\xi_i - 1)\mathcal{J}_i(\tau)\}_{i\in[n]}$ is a sequence of i.n.i.d. random variables. In order to apply the Lindeberg-Feller central limit theorem, we only need to show that (1)

$$\frac{1}{n}\sum_{i\in[n]}\mathcal{J}^2_{n,i}(\tau) \xrightarrow{p} \Sigma(\tau, \tau),$$

where $\Sigma(\tau, \tau)$ is defined in Theorem 3.1, and (2) the Lindeberg condition holds, i.e.,

$$\frac{1}{n}\sum_{i\in[n]}\mathcal{J}^2_{n,i}(\tau)\mathbb{E}(\xi_i - 1)^2 1\{|(\xi_i - 1)\mathcal{J}_{n,i}(\tau)| \geq \sqrt{n}\varepsilon\} \xrightarrow{p} 0.$$

For part (1), we have

$$\frac{1}{n}\sum_{i\in[n]}\mathcal{J}^2_{n,i}(\tau) = \sigma_1^2 + 2\sigma_{12} + \sigma_2^2,$$

where

$$\sigma_1^2 = \frac{1}{n}\sum_{i\in[n]}\left[A_i 1\{S_i = s\}\phi_1(\tau, S_i, Y_i(1), X_i) - (1 - A_i)1\{S_i = s\}\phi_0(\tau, S_i, Y_i(1), X_i)\right]^2,$$

$$\sigma_{12} = \frac{1}{n}\sum_{i\in[n]}\left[A_i 1\{S_i = s\}\phi_1(\tau, S_i, Y_i(1), X_i) - (1 - A_i)1\{S_i = s\}\phi_1(\tau, S_i, Y_i(1), X_i)\right]\phi_s(\tau, S_i),$$

and

$$\sigma_2^2 = \frac{1}{n}\sum_{i\in[n]}\phi_s^2(\tau, S_i).$$

Note

$$
\begin{aligned}
\sigma_1^2 &= \frac{1}{n}\sum_{i\in[n]} A_i 1\{S_i = s\}\phi_1^2(\tau, S_i, Y_i(1), X_i) + \frac{1}{n}\sum_{i\in[n]}(1-A_i)1\{S_i = s\}\phi_0^2(\tau, S_i, Y_i(1), X_i) \\
&\stackrel{d}{=} \frac{1}{n}\sum_{s\in\mathcal{S}} \sum_{i=N(s)+1}^{N(s)+n_1(s)} \phi_1^2(\tau, s, Y_i^s(1), X_i^s) + \frac{1}{n}\sum_{i=N(s)+n_1(s)+1}^{N(s)+n(s)} \phi_0^2(\tau, s, Y_i^s(1), X_i^s) \\
&\xrightarrow{p} \sum_{s\in\mathcal{S}}[\pi(s)\mathbb{E}\phi_1^2(\tau, s, Y_i^s(1), X_i^s) + (1-\pi(s))\mathbb{E}\phi_0^2(\tau, s, Y_i^s(1), X_i^s)] \\
&= \mathbb{E}\left[\pi(S_i)\phi_1^2(\tau, S_i, Y_i(1), X_i) + (1-\pi(s))\phi_0^2(\tau, S_i, Y_i(1), X_i)\right],
\end{aligned}
$$

where the convergence holds due to the fact that $N(s)/n \to F(s)$, $n_1(s)/n \xrightarrow{p} \pi(s)p(s)$, $n(s)/n \xrightarrow{p} p(s)$, and the uniform convergence of the partial sum process. Similarly,

$$
\begin{aligned}
\sigma_{12} &\stackrel{d}{=} \frac{1}{n}\sum_{s\in\mathcal{S}} \sum_{i=N(s)+1}^{N(s)+n_1(s)} \phi_1(\tau, s, Y_i^s(1), X_i^s)\phi_s(\tau, s) + \frac{1}{n}\sum_{i=N(s)+n_1(s)+1}^{N(s)+n(s)} \phi_0(\tau, s, Y_i^s(1), X_i^s)\phi_s(\tau, s) \\
&\xrightarrow{p} \sum_{s\in\mathcal{S}}[\pi(s)\mathbb{E}\phi_1(\tau, s, Y_i^s(1), X_i^s) + (1-\pi(s))\mathbb{E}\phi_0(\tau, s, Y_i^s(1), X_i^s)]\phi_s(\tau, s) = 0,
\end{aligned}
$$

where we use the fact that

$$
\mathbb{E}\phi_1(\tau, s, Y_i^s(1), X_i^s) = \mathbb{E}\phi_0(\tau, s, Y_i^s(1), X_i^s) = 0.
$$

By the standard weak law of large numbers, we have

$$
\sigma_2^2 \xrightarrow{p} \mathbb{E}\phi_s^2(\tau, S_i).
$$

Therefore,

$$
\frac{1}{n}\sum_{i\in[n]} \mathcal{J}_{n,i}^2(\tau) \xrightarrow{p} \mathbb{E}\left[\pi(S_i)\phi_1^2(\tau, S_i, Y_i(1), X_i) + (1-\pi(s))\phi_0^2(\tau, S_i, Y_i(1), X_i)\right] + \mathbb{E}\phi_s^2(\tau, S_i) = \Sigma(\tau, \tau).
$$

To verify the Lindeberg condition, we note that

$$
\begin{aligned}
&\frac{1}{n}\sum_{i\in[n]} \mathcal{J}_{n,i}^2(\tau)\mathbb{E}(\xi_i - 1)^2 1\{|(\xi_i - 1)\mathcal{J}_{n,i}(\tau)| \geq \sqrt{n}\varepsilon\} \\
&\leq \frac{1}{n(\sqrt{n}\varepsilon)^{q-2}}\sum_{i\in[n]} \mathcal{J}_{n,i}^q(\tau)\mathbb{E}(\xi_i - 1)^q \\
&\leq \frac{c}{n(\sqrt{n}\varepsilon)^{q-2}}\sum_{i\in[n]}[\phi_1^q(\tau, S_i, Y_i(1), X_i) + \phi_0^q(\tau, S_i, Y_i(1), X_i) + \phi_s^q(\tau, S_i)] = o_p(1),
\end{aligned}
$$

where the last equality is due to Assumption 3(ii) and the fact that $\eta_{i,a}(s,\tau)$ is bounded.

The finite dimensional convergence of $W_{n,1}^w(\tau) + W_{n,2}^w(\tau)$ across τ can be established in the same manner using the Cramér-Wold device and details are omitted. By the same calculation given above the covariance kernel is shown to be

$$\begin{aligned}
&\lim_{n\to\infty}\frac{1}{n}\sum_{i\in[n]}\mathcal{J}_{n,i}(\tau_1)\mathcal{J}_{n,i}(\tau_2)\\
&=\mathbb{E}\left[\pi(S_i)\phi_1(\tau_1,S_i,Y_i(1),X_i)\phi_1(\tau_2,S_i,Y_i(1),X_i)\right]\\
&+\mathbb{E}\left[(1-\pi(s))\phi_0(\tau_1,S_i,Y_i(1),X_i)\phi(\tau_2,S_i,Y_i(1),X_i)\right]\\
&+\mathbb{E}\phi_s(\tau_1,S_i)\phi_s(\tau_2,S_i)=\Sigma(\tau_1,\tau_2),
\end{aligned}$$

which concludes the proof. □

Lemma P.6. *Suppose Assumptions in Proposition 5.2 hold. Then,*

$$\sup_{\tau\in\Upsilon,s\in\mathcal{S},\theta_1,\theta_0\in\Theta}|Q_n(s,\tau,\theta_1,\theta_0)-Q(s,\tau,\theta_1,\theta_0)|=o_p(1),$$

where $Q(\cdot)$ and $Q_n(\cdot)$ are defined in (5.3) *and* (5.9)*, respectively, with $\Lambda_{a,s}(X_i,\theta_a)=\lambda(\vec{X}_i^\top\theta_a)$. In addition, if $\sup_{\tau\in\Upsilon}(|\hat{f}_1(\hat{q}_1(\tau))-f_1(q_1(\tau))|+|\hat{f}_0(\hat{q}_0(\tau))-f_0(q_0(\tau))|)=o_p(\varepsilon_n)$ for some $\varepsilon_n\downarrow 0$ such that $\varepsilon_n\sqrt{n}\to\infty$. Then,*

$$\sup_{\tau\in\Upsilon,s\in\mathcal{S},\theta_1,\theta_0\in\Theta}|Q_n(s,\tau,\theta_1,\theta_0)-Q(s,\tau,\theta_1,\theta_0)|=o_p(\varepsilon_n).$$

Proof. Recall $\Theta=\Theta_1\times\Theta_0$, where Θ_1 and Θ_0 are two compact sets in $\Re^{d_x}$. Note that

$$\sup_{\theta_a\in\Theta_a}|\hat{g}_{a,s}(X_i,\theta_a)|\le C<\infty,$$

for some constant C. By Assumption 9 and the fact that $\hat{\pi}(s)\xrightarrow{p}\pi(s)>0$, we have

$$Q_n(s,\tau,\theta_1,\theta_0)=\tilde{Q}_n(s,\tau,\theta_1,\theta_0)+R_{n,1}(s,\tau,\theta_1,\theta_0),$$

where

$$\begin{aligned}
\tilde{Q}_n(s,\tau,\theta_1,\theta_0)=&\frac{1}{n(s)}\sum_{i\in I(s)}\left(\frac{\hat{g}_{1,s}(X_i,\theta_1)}{f_1(q_1(\tau))}+\frac{\pi(s)}{1-\pi(s)}\frac{\hat{g}_{0,s}(X_i,\theta_0)}{f_0(q_0(\tau))}\right)^2\\
&+\frac{2}{n_1(s)}\sum_{i\in I_1(s)}\left(\frac{\hat{g}_{1,s}(X_i,\theta_1)}{f_1(q_1(\tau))}+\frac{\pi(s)}{1-\pi(s)}\frac{\hat{g}_{0,s}(X_i,\theta_0)}{f_0(q_0(\tau))}\right)\frac{1\{Y_i\le\hat{q}_1(\tau)\}}{f_1(q_1(\tau))}\\
&+\frac{2\pi(s)}{n_0(s)(1-\pi(s))}\sum_{i\in I_0(s)}\left(\frac{\hat{g}_{1,s}(X_i,\theta_1)}{f_1(q_1(\tau))}+\frac{\pi(s)}{1-\pi(s)}\frac{\hat{g}_{0,s}(X_i,\theta_0)}{f_0(q_0(\tau))}\right)\frac{1\{Y_i\le\hat{q}_0(\tau)\}}{f_0(q_0(\tau))}
\end{aligned}$$

and

$$\sup_{\tau\in\Upsilon,s\in\mathcal{S},\theta_1,\theta_0\in\Theta}|R_{n,1}(s,\tau,\theta_1,\theta_0)|=o_p(1).$$

In addition,

$$\sup_{\theta_1\in\Theta_1}|g_{1,s}(X_i,\theta_1)-\hat{g}_{1,s}(X_i,\theta_1)| = \sup_{\theta_1\in\Theta_1}\left|\frac{1}{n_1(s)}\sum_{i\in I_1(s)}\lambda(\vec{X}_i^\top\theta_1)-\mathbb{E}\lambda(\vec{X}_i^\top\theta_1|S_i=s)\right|$$

$$\stackrel{d}{=}\sup_{\theta_1\in\Theta_1}\left|\frac{1}{n_1(s)}\sum_{i=N(s)+1}^{N(s)+n_1(s)}(\lambda((\vec{X}_i^s)^\top\theta_1)-\mathbb{E}\lambda((\vec{X}_i^s)^\top\theta_1))\right| = O_p(n^{-1/2}), \tag{P.4}$$

where the last equality is due to the fact that $\{X_i^s\}_{i\in[n]}$ is an i.i.d. sequence independent of $\{A_i,S_i\}_{i\in[n]}$ and the usual maximal inequality such as that in van der Vaart and Wellner (1996, Theorem 2.14.1) applies. Therefore,

$$\tilde{Q}_n(s,\tau,\theta_1,\theta_0)=\breve{Q}_n(s,\tau,\theta_1,\theta_0)+R_{n,2}(s,\tau,\theta_1,\theta_0),$$

where

$$\begin{aligned}
\breve{Q}_n(s,\tau,\theta_1,\theta_0) &= \frac{1}{n(s)}\sum_{i\in I(s)}\left(\frac{g_{1,s}(X_i,\theta_1)}{f_1(q_1(\tau))}+\frac{\pi(s)}{1-\pi(s)}\frac{g_{0,s}(X_i,\theta_0)}{f_0(q_0(\tau))}\right)^2 \\
&+\frac{2}{n_1(s)}\sum_{i\in I_1(s)}\left(\frac{g_{1,s}(X_i,\theta_1)}{f_1(q_1(\tau))}+\frac{\pi(s)}{1-\pi(s)}\frac{g_{0,s}(X_i,\theta_0)}{f_0(q_0(\tau))}\right)\frac{1\{Y_i\le\hat{q}_1(\tau)\}}{f_1(q_1(\tau))} \\
&+\frac{2\pi(s)}{n_0(s)(1-\pi(s))}\sum_{i\in I_0(s)}\left(\frac{g_{1,s}(X_i,\theta_1)}{f_1(q_1(\tau))}+\frac{\pi(s)}{1-\pi(s)}\frac{g_{0,s}(X_i,\theta_0)}{f_0(q_0(\tau))}\right)\frac{1\{Y_i\le\hat{q}_0(\tau)\}}{f_0(q_0(\tau))} \\
&\equiv \breve{Q}_{n,1}(s,\tau,\theta_1,\theta_0)+\breve{Q}_{n,2}(s,\tau,\theta_1,\theta_0)+\breve{Q}_{n,3}(s,\tau,\theta_1,\theta_0)
\end{aligned}$$

and

$$\sup_{\tau\in\Upsilon,s\in\mathcal{S},\theta_1,\theta_0\in\Theta}|R_{n,2}(s,\tau,\theta_1,\theta_0)|=O_p(n^{-1/2}).$$

By the same argument as (P.4), we can show that

$$\sup_{\tau\in\Upsilon,s\in\mathcal{S},\theta_1,\theta_0\in\Theta}\left|\breve{Q}_{n,1}(s,\tau,\theta_1,\theta_0)-\mathbb{E}\left[\left(\frac{g_{1,s}(X_i,\theta_1)}{f_1(q_1(\tau))}+\frac{\pi(s)}{1-\pi(s)}\frac{g_{0,s}(X_i,\theta_0)}{f_0(q_0(\tau))}\right)^2\middle|S_i=s\right]\right|=O_p(n^{-1/2}).$$

Next, we show

$$\sup_{\tau\in\Upsilon,s\in\mathcal{S},\theta_1,\theta_0\in\Theta}\left|\breve{Q}_{n,2}(s,\tau,\theta_1,\theta_0)-2\mathbb{E}\left[\left(\frac{g_{1,s}(X_i,\theta_1)}{f_1(q_1(\tau))}+\frac{\pi(s)}{1-\pi(s)}\frac{g_{0,s}(X_i,\theta_0)}{f_0(q_0(\tau))}\right)\frac{1\{Y_i(1)\le q_1(\tau)\}}{f_1(q_1(\tau))}|S_i=s\right]\right|$$
$$=O_p(n^{-1/2}).$$

The proof of

$$\sup_{\tau\in\Upsilon,s\in\mathcal{S},\theta_1,\theta_0\in\Theta}\left|\breve{Q}_{n,3}(s,\tau,\theta_1,\theta_0)-\frac{2\pi(s)}{(1-\pi(s))}\mathbb{E}\left[\left(\frac{g_{1,s}(X_i,\theta_1)}{f_1(q_1(\tau))}+\frac{\pi(s)}{1-\pi(s)}\frac{g_{0,s}(X_i,\theta_0)}{f_0(q_0(\tau))}\right)\frac{1\{Y_i\le q_0(\tau)\}}{f_0(q_0(\tau))}\middle|S_i=s\right]\right|$$

$$= O_p(n^{-1/2})$$

is similar and is omitted.

Denote

$$\phi_i(s,\tau,\theta_1,\theta_0,q) = \left(\frac{g_{1,s}(X_i,\theta_1)}{f_1(q_1(\tau))} + \frac{\pi(s)}{1-\pi(s)}\frac{g_{0,s}(X_i,\theta_0)}{f_0(q_0(\tau))}\right)\frac{1\{Y_i(1)\leq q\}}{f_1(q_1(\tau))}.$$

Because $\sup_{\tau\in\Upsilon}|\hat{q}_1(\tau)-q_1(\tau)| = O_p(n^{-1/2})$, for any $\varepsilon > 0$, there exists a constant $M > 0$ such that with probability greater than $1-\varepsilon$, we have

$$\sup_{\tau\in\Upsilon}|\hat{q}_1(\tau)-q_1(\tau)| \leq n^{-1/2}M.$$

Therefore, with probability greater than $1-\varepsilon$, we have

$$\begin{aligned}
&\sup_{\tau\in\Upsilon,s\in\mathcal{S},\theta_1,\theta_0\in\Theta}\left|\breve{Q}_{n,2}(s,\tau,\theta_1,\theta_0) - 2\mathbb{E}\left[\left(\frac{g_{1,s}(X_i,\theta_1)}{f_1(q_1(\tau))} + \frac{\pi(s)}{1-\pi(s)}\frac{g_{0,s}(X_i,\theta_0)}{f_0(q_0(\tau))}\right)\frac{1\{Y_i(1)\leq q_1(\tau)\}}{f_1(q_1(\tau))}|S_i=s\right]\right| \\
&\leq \sup_{\tau\in\Upsilon,s\in\mathcal{S},\theta_1,\theta_0\in\Theta,|q_1-q_2|\leq n^{-1/2}M}\left|\frac{2}{n_1(s)}\sum_{i\in I_1(s)}\left[\phi_i(s,\tau,\theta_1,\theta_0,q_1) - \mathbb{E}(\phi_i(s,\tau,\theta_1,\theta_0,q_2)|S_i=s)\right]\right| \\
&\leq \sup_{\tau\in\Upsilon,s\in\mathcal{S},\theta_1,\theta_0\in\Theta,q_1\in\Re}\left|\frac{2}{n_1(s)}\sum_{i\in I_1(s)}\left[\phi_i(s,\tau,\theta_1,\theta_0,q_1) - \mathbb{E}(\phi_i(s,\tau,\theta_1,\theta_0,q_1)|S_i=s)\right]\right| \\
&\quad + \sup_{\tau\in\Upsilon,s\in\mathcal{S},\theta_1,\theta_0\in\Theta,|q_1-q_2|\leq n^{-1/2}M}\left|2\left[\mathbb{E}(\phi_i(s,\tau,\theta_1,\theta_0,q_1)|S_i=s) - \mathbb{E}(\phi_i(s,\tau,\theta_1,\theta_0,q_2)|S_i=s)\right]\right|.
\end{aligned} \tag{P.5}$$

For the first term on the RHS of (P.5), we note that

$$\mathcal{F} = \{\phi_i(s,\tau,\theta_1,\theta_0,q_1) : \tau\in\Upsilon, s\in\mathcal{S},\theta_1,\theta_0\in\Theta, q_1\in\Re\}$$

is of the VC-type with fixed coefficients (α, v) and a bounded envelope. Therefore, Lemma P.2 implies

$$\sup_{\tau\in\Upsilon,s\in\mathcal{S},\theta_1,\theta_0\in\Theta,q_1\in\Re}\left|\frac{2}{n_1(s)}\sum_{i\in I_1(s)}\left[\phi_i(s,\tau,\theta_1,\theta_0,q_1) - \mathbb{E}(\phi_i(s,\tau,\theta_1,\theta_0,q_1)|S_i=s)\right]\right| = O_p(n^{-1/2}).$$

For the second term on the RHS of (P.5), we note that for some constant $C > 0$,

$$\sup_{\tau\in\Upsilon,s\in\mathcal{S},\theta_1,\theta_0\in\Theta,q_1\in\Re}|\partial_q\mathbb{E}(\phi_i(s,\tau,\theta_1,\theta_0,q)|S_i=s)| \leq C\sup_{q\in\Re}f_1(q|X_i,S_i) < \infty.$$

Therefore,

$$\sup_{\tau\in\Upsilon,s\in\mathcal{S},\theta_1,\theta_0\in\Theta,|q_1-q_2|\leq n^{-1/2}M}\left|2\left[\mathbb{E}(\phi_i(s,\tau,\theta_1,\theta_0,q_1)|S_i=s) - \mathbb{E}(\phi_i(s,\tau,\theta_1,\theta_0,q_2)|S_i=s)\right]\right| = O(n^{-1/2}).$$

Combining these bounds, we have shown that

$$\sup_{\tau\in\Upsilon,s\in\mathcal{S},\theta_1,\theta_0\in\Theta}|Q_n(s,\tau,\theta_1,\theta_0)-Q(s,\tau,\theta_1,\theta_0)|=o_p(1).$$

For the second result, we note that

$$\sup_{\tau\in\Upsilon,s\in\mathcal{S},\theta_1,\theta_0\in\Theta}|R_{n,1}(s,\tau,\theta_1,\theta_0)|\leq C\sup_{\tau\in\Upsilon}(|\hat{f}_1(\hat{q}_1(\tau))-f_1(q_1(\tau))|+|\hat{f}_0(\hat{q}_0(\tau))-f_0(q_0(\tau))|)=o_p(\varepsilon_n).$$

The other terms converge at the $n^{1/2}$-rate. This implies

$$\sup_{\tau\in\Upsilon,s\in\mathcal{S},\theta_1,\theta_0\in\Theta}|Q_n(s,\tau,\theta_1,\theta_0)-Q(s,\tau,\theta_1,\theta_0)|=o_p(\varepsilon_n).$$

□

Lemma P.7. *Suppose Assumptions in Theorem 5.4 hold. Then,*

$$\sup_{\tau\in\Upsilon,a=0,1,s\in\mathcal{S}}\left\|\frac{1}{n_a(s)}\sum_{i\in I_a(s)}(1\{Y_i\leq\hat{q}_a(\tau)\}-m_a(\tau,s,X_i))H_{h_n}(X_i)\right\|_\infty=O_p\left(\sqrt{\frac{\log(n)}{n}}\right).$$

Proof. We focus on $a=1$. We have

$$\begin{aligned}
&\sup_{\tau\in\Upsilon}\left\|\frac{1}{n_1(s)}\sum_{i\in I_1(s)}(1\{Y_i\leq\hat{q}_1(\tau)\}-m_1(\tau,s,X_i))H_{h_n}(X_i)\right\|_\infty\\
&\leq\sup_{\tau\in\Upsilon}\left\|\frac{1}{n_1(s)}\sum_{i\in I_1(s)}(1\{Y_i\leq\hat{q}_1(\tau)\}-\mathbb{P}(Y_i(1)\leq\hat{q}_1(\tau)|X_i,S_i=s))H_{h_n}(X_i)\right\|_\infty\\
&+\sup_{\tau\in\Upsilon}\left\|\frac{1}{n_1(s)}\sum_{i\in I_1(s)}(\mathbb{P}(Y_i(1)\leq\hat{q}_1(\tau)|X_i,S_i=s)-m_1(\tau,s,X_i))H_{h_n}(X_i)\right\|_\infty\\
&\leq\sup_{q\in\Re}\left\|\frac{1}{n_1(s)}\sum_{i\in I_1(s)}(1\{Y_i\leq q\}-\mathbb{P}(Y_i(1)\leq q|X_i,S_i=s))H_{h_n}(X_i)\right\|_\infty\\
&+\sup_{\tau\in\Upsilon}\left\|\frac{1}{n_1(s)}\sum_{i\in I_1(s)}(\mathbb{P}(Y_i(1)\leq\hat{q}_1(\tau)|X_i,S_i=s)-m_1(\tau,s,X_i))H_{h_n}(X_i)\right\|_\infty.
\end{aligned}\tag{P.6}$$

Denote

$$\mathcal{F}_h=\{1\{Y_i^s(1)\leq q\}H_{h_n,h}(X_i):q\in\Re\},\quad\mathcal{F}=\cup_{h\in[h_n]}\mathcal{F}_{h_n},$$

and $H_{h_n,h}(X_i)$ is the h-th coordinate of $H_{h_n}(X_i)$. For each $h\in[h_n]$, $\mathcal{F}_h$ is of the VC-type with fixed coefficients (α, v) and a common envelope $F_i = ||H_{h_n}(X_i)||_2 \leq \zeta(h_n)$, i.e.,

$$\sup_Q N(\mathcal{F}_h, e_Q, \varepsilon||F||_{Q,2}) \leq \left(\frac{\alpha}{\varepsilon}\right)^v, \quad \forall \varepsilon \in (0,1],$$

where the supremum is taken over all finitely discrete probability measures. This implies

$$\sup_Q N(\mathcal{F}, e_Q, \varepsilon||F||_{Q,2}) \leq \sum_{h \in [h_n]} \sup_Q N(\mathcal{F}_h, e_Q, \varepsilon||F||_{Q,2}) \leq \left(\frac{\alpha h_n}{\varepsilon}\right)^v \quad \forall \varepsilon \in (0,1],$$

i.e., $\mathcal{F}$ is also of the VC-type with coefficients $(\alpha h_n, v)$. In addition,

$$\sup_{f \in \mathcal{F}} \mathbb{E} f^2 \leq \max_{h \in [h_n]} \mathbb{E} H^2_{h_n,h}(X_i) \leq C < \infty.$$

Then, Lemma P.2 implies

$$\begin{aligned}
&\sup_{q \in \Re} \left\| \frac{1}{n_1(s)} \sum_{i \in I_1(s)} (1\{Y_i \leq q\} - \mathbb{P}(Y_i(1) \leq q | X_i, S_i = s)) H_{h_n}(X_i) \right\|_\infty \\
&= O_p\left(\sqrt{\frac{\log(h_n \zeta(h_n))}{n}} + \frac{\zeta(h_n)\log(\zeta(h_n))}{n}\right) = O_p\left(\sqrt{\frac{\log(n)}{n}}\right).
\end{aligned}$$

For the second term of (P.6), because $\sup_{q \in \Re, x \in \text{Supp}(X), s \in \mathcal{S}} f_1(q|x,s) < \infty$, we have

$$\begin{aligned}
&\sup_{\tau \in \Upsilon} \left\| \frac{1}{n_1(s)} \sum_{i \in I_1(s)} (\mathbb{P}(Y_i(1) \leq \hat{q}_1(\tau) | X_i, S_i = s) - m_1(\tau, s, X_i)) H_{h_n}(X_i) \right\|_\infty \\
&\leq \sup_{\tau \in \Upsilon} |\hat{q}_1(\tau) - q_1(\tau)| \left\| \frac{1}{n_1(s)} \sum_{i \in I_1(s)} |H_{h_n}(X_i)| \right\|_\infty \\
&\leq \sup_{\tau \in \Upsilon} |\hat{q}_1(\tau) - q_1(\tau)| \left\| \frac{1}{n_1(s)} \sum_{i \in I_s(1)} [|H_{h_n}(X_i^s)| - \mathbb{E}(|H_{h_n}(X_i)||S_i = s)] \right\|_\infty \\
&+ \sup_{\tau \in \Upsilon} |\hat{q}_1(\tau) - q_1(\tau)| \, \|\mathbb{E}(|H_{h_n}(X_i)||S_i = s)\|_\infty \\
&= \sup_{\tau \in \Upsilon} |\hat{q}_1(\tau) - q_1(\tau)| \left\| \frac{1}{n_1(s)} \sum_{i \in I_s(1)} [|H_{h_n}(X_i^s)| - \mathbb{E}(|H_{h_n}(X_i)||S_i = s)] \right\|_\infty + O_p(n^{-1/2}) \\
&= O_p(n^{-1/2}),
\end{aligned}$$

where the second to last inequality holds because of Assumption 7 and $\|\mathbb{E}(|H_{h_n}(X_i)||S_i = s)\|_\infty \leq C < \infty$, and the last inequality holds because by a similar argument to the one used in bounding the first term on the RHS of (P.6), we can show that

$$\left\| \frac{1}{n_1(s)} \sum_{i \in I_s(1)} [|H_{h_n}(X_i^s)| - \mathbb{E}(|H_{h_n}(X_i)||S_i = s)] \right\|_\infty = O_p\left(\sqrt{\frac{\log(n)}{n}}\right).$$

This concludes the proof. $\square$

Lemma P.8. *Suppose Assumptions in Theorem 5.4 hold and recall* $\overline{\ell}$ *defined in* (M.1). *We have* $\overline{\ell}/(\sqrt{h_n \log(n)}/n) \to \infty$*, w.p.a.1.*

Proof. Note that w.p.a.1,

$$\begin{aligned} \overline{\ell} &= \inf_{U \in \Re^{h_n}} \frac{\left[\frac{1}{n_a(s)} \sum_{i \in I_a(s)} (H_{h_n}^\top(X_i)U)^2\right]^{3/2}}{\frac{1}{n_a(s)} \sum_{i \in I_a(s)} |H_{h_n}^\top(X_i)U|^3} \\ &\geq \inf_{U \in \Re^{h_n}} \frac{\left[\frac{1}{n_a(s)} \sum_{i \in I_a(s)} (H_{h_n}^\top(X_i)U)^2\right]^{1/2}}{\sup_{x \in \mathcal{X}} ||H_{h_n}(x)||_2 ||U||_2} \geq \frac{\kappa_1^{1/2}}{\zeta(h_n)}, \end{aligned}$$

where the last inequality is due to Assumption 11. Therefore,

$$\overline{\ell}/(\sqrt{h_n \log(n)}/n) \geq \sqrt{\frac{\kappa_1 n}{\zeta^2(h_n) h_n \log(n)}} \to \infty \ w.p.a.1.$$

$\square$

References

Bach, F. (2010). Self-concordant analysis for logistic regression. *Electronic Journal of Statistics 4*, 384–414.

Belloni, A., V. Chernozhukov, I. Fernández-Val, and C. Hansen (2017). Program evaluation with high-dimensional data. *Econometrica 85*(1), 233–298.

Bugni, F. A., I. A. Canay, and A. M. Shaikh (2018). Inference under covariate-adaptive randomization. *Journal of the American Statistical Association 113*(524), 1741–1768.

Chernozhukov, V., D. Chetverikov, and K. Kato (2014). Gaussian approximation of suprema of empirical processes. *The Annals of Statistics 42*(4), 1564–1597.

Kato, K. (2009). Asymptotics for argmin processes: Convexity arguments. *Journal of Multivariate Analysis 100*(8), 1816–1829.

Knight, K. (1998). Limiting distributions for l1 regression estimators under general conditions. *The Annals of Statistics 26*(2), 755–770.

Montgomery-Smith, S. J. (1993). Comparison of sums of independent identically distributed random variables. *Probability and Mathematical Statistics 14* (2), 281–285.

van der Vaart, A. and J. A. Wellner (1996). *Weak Convergence and Empirical Processes*. Springer, New York.

Zhang, Y. and X. Zheng (2020). Quantile treatment effects and bootstrap inference under covariate-adaptive randomization. *Quantitative Economics 11* (3), 957–982.